ATS Method : Periodogram Smoothing by Non-Parametric Regression

by

NARASIMHA G. PAI

DEPARTMENT OF ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY KANPUR

MARCH, 1995

ATS Method : Periodogram Smoothing by Non-Parametric Regression

A Thesis Submitted in Partial
Fulfilment of the Requirements
for the Degree of
Master of Technology
by

Narasimha G. Pai

Department of Electrical Engineering
Indian Institute of Technology Kanpur

March 1995

CERTIFICATE

It is certified that the work contained in the thesis titled ATS Method : Periodogram Smoothing by Non-Parametric Regression, by *Narasimha G. Pai* has been carried out under my supervision and that this work has not been submitted elsewhere for a degree.

March, 1995

Dr. Govind Sharma,
Associate Professor,
Department of Electrical Engineering
IIT, Kanpur.

Dedicated to

My Parents

ACKNOWLEDGEMENTS

I would like to express my gratitude to Dr.Govind Sharma for giving me the opportunity to work under him.

I would like to thank all my M.tech batchmates for the memorable moments that we all had together. Also I would specially like to thank Amit, Kartik and Bhavesh for helping me out of difficulties during the preparation of the thesis report. I will always remember and cherish the good times I had playing cricket and football with my inmates of Hall-4. I shall also remember the workers in our hall canteen (especially chotu) for providing me with tea and eats with an eversmiling attitude even at odd hours. Lastly I must mention the members of G-5 (Bhavesh, Sanjay, Varda and Venu) for the endless hours of thulla we had together.

My stay here was possible due to the constant moral support provided by my family members.

March 29, 1995 Narasimha G. Pai

Abstract

Many well known techniques exist for smoothing the periodogram estimate of power spectrum viz Bartlett's and Windowing methods. ATS is a three step procedure for the same. First a small amount of local averaging ('A') is carried out on the periodogram. Then a variance stabilizing transform is applied ('T') and finally the result is smoothed ('S') using a non-parametric regression procedure. This method works well even when data distributions are not necessarily Gaussian. Simulations were carried out on different data distributions using ATS and Bartlett's method. Simulation results show that in a number of cases the proposed ATS method is superior to the Bartlett's method.

Contents

1 Introduction **1**

1.1 Preliminaries 1

1.2 Spectrum Estimation Methods 2

1.3 Thesis Organisation 3

2 Review **5**

2.1 Periodogram estimate of power spectrum 5

2.1.1 Bias of periodogram estimate 6

2.1.2 Variance of the periodogram estimate 6

2.2 Smoothed Spectrum Estimators 7

2.2.1 Bartlett's method of averaging periodogram 8

2.2.2 Windowing method 9

3 The ATS method **11**

3.1 Basic Theory 11

3.2 The Spectral Estimation problem 12

3.2.1 Distribution of the estimate 12

3.3 The averaging step 'A' 14

3.3.1 Bias and variance of the averaged estimates 14

3.4 The variance stabilizing transform 'S' 15

3.5 The smoothing process 'S' 17

3.6 Choice of various parameters in the loess algorithm 19

3.7 Scaled Exponential Variables 21

4 Simulation Results and Conclusions **23**

4.1 Generation of the Random Signal 23

4.2 Simulations 24

4.2.1 Simulations for the Stabilizing Transform 24

4.2.2 Simulations for Periodogram Smoothing 25

4.3 Conclusions 29

5 Future Scope **42**

List of Figures

4.1 Power Spectrum of the AR2 Process y[n] 26

4.2 Estimate Variance vs Frequency.(a=2) 27

4.3 Estimate Variance vs Frequency.(a=4) 28

4.4 Exponential White Noise Spectrum Estimate by ATS and Bartlett's Methods.(a=2) 31

4.5 Exponential White Noise Spectrum Estimate by ATS and Bartlett's Method.(a=4) 32

4.6 Exponential White Noise Spectrum Estimate by ATS and Bartlett's Method.(a=8) 33

4.7 Spectrum Estimate of a Narrow Band Signal by ATS Method.(a=2) . 34

4.8 Spectrum Estimate of a Narrow Band Signal by Bartlett's Method. (Average of Two) 35

4.9 Spectrum Estimate of a Narrow Band Signal by ATS Method.(a=4) . 36

4.10 Spectrum Estimate of a Narrow Band Signal by Bartlett's Method. (Average of Four) 37

4.11 Spectrum Estimate of a Narrow Band Signal by ATS Method.(a=8) . 38

4.12 Spectrum Estimate of A Narrow Band Signal by Bartlett's Method. (Average of Eight) 39

4.13 Spectrum Estimate of Sinusoids in Uniform White Noise by ATS Method.(a=2) 40

4.14 Spectrum Estimate of Sinusoids in Uniform White Noise by Bartlett's Method.(Average of Two) 41

Chapter 1

Introduction

1.1 Preliminaries

The need for power spectrum arises in a variety of contexts in communication systems viz measurement of noise for design of optimal linear filter, measurement of narrowband signal in wideband noise, estimation of system parameters using an input noise excitation and measuring the output. Starting from the problem of estimation of power spectrum we are often reduced to the problem of necessity for accurate measurement and fast computations. For exact measurement of power we would require an accurately measured infinite record length of the random signal which in turn necessitates infinite amounts of computation. These requirements being impractical an approximate measurement is necessary. Such an approximate measurement of power spectrum raises the issue of what record length of the random signal is required for a given accuracy, what computational approach should be used etc. Practical answers to these questions can be obtained by using certain results of statistical estimation theory.

In our discussion we shall use two properties of spectral estimators viz bias and variance. Generally speaking in case of parameter estimation if $\hat{\alpha}$ is an estimate of a parameter α then

$$Bias = \alpha - \varepsilon[\hat{\alpha}]$$

$$Variance = \varepsilon(\hat{\alpha} - \varepsilon[\hat{\alpha}])^2$$

An unbiased estimator is one for which bias is zero. This means that the expected value of the estimator is equal to the true value of the parameter to be measured. A small value of variance suggests that the probability distribution function of the estimator is concentrated around its mean(true value in case of an unbiased estimator). If for an estimator the bias and variance both tend to zero as the record length increases then such an estimator is called a consistent estimator [16].

Communication Systems are generally required to handle a variety of signals in the presence of noise. To a large extent the design of such systems depends on the statistical properties of both the signal and noise. In most cases the noise can be represented (or approximated) as stationary Gaussian random process with zero mean. Certain applications require that signals and noises which are approximately stationary but not Gaussian be studied for their autocovariance or power spectrum. Although in a non-Gaussian framework autocovariance and spectrum are no longer the only relevent statistics, they are usually the most important ones [4]. However in either case measurement of power is of interest.

We have only considered the case of measurement of spectra of individual signals or noises. The measurement of cross-spectrum is not discussed here.

1.2 Spectrum Estimation Methods

Broadly there are two methods for estimation of power spectrum.

- Classical power spectrum estimation based on the Fourier analysis of the measured random signal.
- Parametric power spectrum estimation wherein a statistical model is chosen(AR, MA or ARMA) for the random signal and further analysis is carried out.

The classical method applies the Fourier Transform on the random signal. The magnitude of the Fourier Transform squared divided by the record length is then a estimate of power. This is actually the same as the Fourier Transform of the

ACF of the random signal. Since we are considering only a finite record length of the random signal we are actually working with a windowed signal and a windowed ACF. Windowing means that we are assuming the signal and ACF to be zero outside the chosen window(The window size depends on the chosen record length). Often we have more knowledge about the random signal so that a more reasonable assumption than the signal or ACF as zero outside the chosen record length can be made [14]. This is done by choosing an appropriate model for the random signal. The problem then reduces to the estimation of the parameters of this model. This parametric spectrum estimation method is not used in our work.

As shall be seen in chapter 2 the classical method does not give a consistent estimate. To make it a consistent estimate and to smooth the spectrum there exist well known techniques viz the Bartlett and Windowing methods. This work investigates the application of a new method (ATS method) to the same problem [8]. The estimate obtained from this method is compared with the one obtained with the Bartlett's method. Since the method is supposed to work for non-Gaussian signals as well, simulations are carried out on non-Gaussian signals.

1.3 Thesis Organisation

The organisation of the remaining chapters is as given

- As mentioned previously chapter 2 reviews the existing methods of periodogram smoothing viz Bartlett and Windowing method. These estimators are for their bias and variance.

- Chapter 3 introduces the ATS method. This chapter describes in detail the method and the various steps involved.

- Chapter 4 contains the simulations carried out using the ATS and Bartlett's method. Comparisons between the two methods are made from the plots of the estimated power spectrum. The conclusions based on these plots are also listed in this chapter.

- Chapter 5 discusses the scope for further work using this method and improvements are suggested.

Chapter 2

Review

In this chapter we will review the existing methods of smoothing the periodogram spectrum estimate. We will compare the advantages and disadvantages of these methods in terms of their bias and variance.

2.1 Periodogram estimate of power spectrum

For estimation of ACF and power spectrum of a random signal using time averages the signal should be ergodic. For an ergodic signal the time averages will approach the statistical average for increasing record length [17]. For any random signal there are two generally used estimators of ACF; one providing an unbiased estimate while the other a biased one. Both however provide a consistent asymtotically unbiased estimate. Unfortunately the Fourier Transform of this estimate is not a consistent estimate of the power spectrum. This is because the variance of the estimate does not tend to zero for increasing record length N of the random process. However smoothing this estimate does provide a consistent estimate of the power spectrum. For a discrete time real random process $x[n]$ let $c_{xx}[m]$ be the estimate of the autocovariance and $I_N(\omega)$ be the estimate of its power spectrum. Then for length $2M+1$ of the ACF $c_{xx}[m]$ we have

$$I_N(\omega) = \sum_{m=1}^{M} c_{xx}[m] \exp^{-j\omega n} \tag{2.1}$$

Corresponding to this if $x[n]$ has a length N and Fourier Transform

$$X(\exp^{j\omega}) = \sum_{n=0}^{N} x[n] \exp^{-j\omega n} \tag{2.2}$$

it can be shown that [16] eqn 2.1 is the same as

$$I_N(\omega) = \frac{1}{N} \mid X(\exp^{j\omega})^2 \mid \tag{2.3}$$

This is known as the periodogram estimate of power.

2.1.1 Bias of periodogram estimate

From eqns 2.1 and 2.3 we see that the periodogram estimate is equivalent to the Fourier Transform of the biased autocovariance estimator $c_{xx}[m]$. Taking the expectation of eq 2.1 we have

$$\varepsilon[I_N(\omega)] = \sum_{m=1}^{M} \varepsilon[c_{xx}[m]] \exp^{-j\omega n}$$

But from [16]

$$\varepsilon[c_{xx}[m]] = \frac{N - \mid m \mid}{N} \phi_{xx}(m) \quad \mid m \mid < M$$

Here $\phi_{xx}(m)$ is the autocovariance of the random process $x[n]$. Due to the finite limits of summation and the factor $\frac{N-|m|}{N}$ the expectation of $I_N(\omega)$ is not equal to the Fourier Transform of $\phi_{xx}(\omega)$. Therefore the periodogram is a biased estimator of the actual power spectrum $P_{xx}(\omega)$. Thus $\varepsilon[I_N(\omega)]$ can be viewed as the Fourier Transform of the product of $\frac{N-|m|}{N}$ and $\phi_{xx}(m)$. The equivalent frequency domain representation is

$$\varepsilon[I_N(\omega)] = \frac{1}{2\pi} \int_{-\pi}^{\pi} P_{xx}(\theta) W_B(\exp^{j(\omega-\theta)}) d\theta \tag{2.4}$$

where $W_B(\exp^{j\omega}) = \frac{1}{N}(\frac{sin(\omega N/2)}{sin(\omega/2)})^2$ is the Fourier Transform of $\frac{N-|m|}{N}$.

2.1.2 Variance of the periodogram estimate

We first consider the case when $x[n]$ is a white Guassian noise sequence. For this case it can be shown that [10]

$$\varepsilon(x[k]x[l]x[m]x[n]) \;=\; \sigma_x^4 \; k = l \; and \; m = n \; or$$

$$k = m \; and \; l = n \; or$$

$$k = n \; and \; l = m$$

$$= \quad 0 \; otherwise$$

Then the covariance of the estimate is given by [16]

$$Cov[I_N(\omega_1), I_N(\omega_2)] = \sigma_x^4[(\frac{sin[(\omega_1+\omega_2)N/2]}{Nsin[(\omega_1+\omega_2)/2]})^2 + (\frac{sin[(\omega_1-\omega_2)N/2]}{Nsin[(\omega_1-\omega_2)/2]})^2] \quad (2.5)$$

Substituting $\omega_1 = \omega_2$ in eq 2.5 we get the variance of the periodogram in case of white gaussian noise as

$$Var[I_N(\omega)] = \sigma^4[1 + (\frac{sin[\omega N]}{Nsin[\omega]})^2] \quad (2.6)$$

From eqn 2.6 we see that $Var[I_N(\omega)]$ is of the order of σ^4. Also the variance does not approach zero as N increases. Thus the periodogram is not a consistent estimate of the power spectrum. From eqn 2.5 we see that the periodogram estimates seperated by integer multiples of $\frac{2\pi}{N}$ are uncorrelated. From this property and the non-consistency of the estimate we can conclude that as N increases the fluctuations in the periodogram become rapid.

For the case where $x[n]$ is a coloured noise we assume that it is derived from a white noise with unit variance by passing it through a linear filter. The magnitude squared response of this filter is $P_{xx}(\omega)$, where $P_{xx}(\omega)$ is the true value of the power spectrum of coloured noise. Let $x_N[n]$ be a N point sequence of the coloured noise. If $I_N(\omega)$ is the estimate of periodogram for $x_N[n]$ then [13]

$$Cov[I_N(\omega_1), I_N(\omega_2)] = P_{xx}(\omega_1)P_{xx}(\omega_2)[(\frac{sin[(\omega_1+\omega_2)N/2]}{Nsin[(\omega_1+\omega_2)/2]})^2 + (\frac{sin[(\omega_1-\omega_2)N/2]}{Nsin[(\omega_1-\omega_2)/2]})^2] \quad (2.7)$$

Here too we see that the frequencies which are integer multiples of $\frac{2\pi}{N}$ apart are uncorrelated. Also the estimate as in the case of white noise is not a consistent estimate.

2.2 Smoothed Spectrum Estimators

We shall now see two techniques to obtain a consistent estimate from the periodogram viz Bartlett and Windowing methods.

2.2.1 Bartlett's method of averaging periodogram

In this method a number of independent estimates are taken and averaged to obtain a reduction in the variance of the estimate. The origin of this method is attributed to Bartlett; hence the name.

For a random process x[n] a sequence of length N is divided into K segments of length M each such that $N = KM$. Therefore

$$x^{(i)}[n] = x[n + iM - M] \;\; 0 <= n <= M - 1; \;\; 1 <= i <= K$$

The K periodogram estimates are computed as

$$I_M^{(i)}(\omega) = \frac{1}{M} \mid \sum_{n=0}^{M-1} x^{(i)}[n] \exp^{-j\omega n} \mid^2$$

If the values of the autocovariance function $\phi_{xx}[m]$ is small for $m > M$ then we can assume that the estimates of $I_M^{(i)}$ are independent of each other. Then the Bartlett's estimate is defined as

$$B_{xx}(\omega) = \frac{1}{K} \sum_{i=1}^{K} I_M^{(i)}(\omega)$$

$$\varepsilon[B_{xx}(\omega)] = \frac{1}{K} \sum_{i=1}^{K} \varepsilon[I_M^{(i)}] = \varepsilon[I_M^{(i)}]$$

Substituting eqn 2.4 of sec 2.1.1 for $\varepsilon[I_M^{(i)}]$ we have

$$\varepsilon[B_{xx}(\omega)] = \frac{1}{2\pi M} \int_{-\pi}^{\pi} P_{xx}(\theta) (\frac{sin[(\omega - \theta)M/2]}{sin[(\omega - \theta)/2]})^2 d\theta$$

Thus the expected value of the Bartlett estimate is equal to the convolution of the true spectrum with the Fourier Transform of the triangular window function corresponding to the M sample periodogram. Thus the Bartlett estimate is also a biased estimate of the spectrum. However since the estimate is obtained by averaging independent estimates, the variance is

$$Var[B_{xx}(\omega)] = \frac{1}{K} Var[I_M(\omega)] = \frac{1}{K} [P_{xx}(\omega)^2][1 + (\frac{sin(\omega M)}{M sin(\omega)})^2] \qquad (2.8)$$

As the number of independent estimates increases the variance reduces and starts approaching zero. Thus the Bartlett estimate is a consistent estimate of power.

We have seen that the bias of the estimate is the convolution of the true spectrum with a spectral window. The width of the main lobe of the spectral window is inversely proportional to M. The bias increases with the increased width of the main lobe, consequently the spectral resolution reduces. For a fixed recod length as the number of periodograms increases the variance of the estimate reduces but due to a decrease in M the spectral resolution also reduces. Thus we see that there is a trade-off between the bias and the variance.

2.2.2 Windowing method

An alternative approach to smoothing the periodogram is by convolution with an appropriate spectral window. If $S_{xx}(\omega)$ is the smoothed estimate and W is the appropriate spectral window then

$$S_{xx}(\omega) = \frac{1}{2\pi}\int_{-\pi}^{\pi} I_N(\theta) W(\exp^{j(\omega\theta)}) d\theta$$

$$If \quad w[m] = \frac{1}{2\pi}\int_{-\pi}^{\pi} W(\exp^{j\omega}) \exp^{j\omega m} d\omega$$

is a finite sequence of length $2M-1$ then

$$S_{xx}(\omega) = \sum_{m=-(M-1)}^{(M-1)} c_{xx}(m) w(m) \exp^{-j\omega m}$$

Here $c_{xx}[m]$ is the estimate of autocovariance. w[m] should be an even sequence so that $S_{xx}(\omega)$ is even and real when the input sequence $x[n]$ is real. Also since the power spectrum is a non-negative quantity a sufficient although not necessary condition to ensure that $S_{xx}(\omega)$ is non-negative is $W(\exp^{j\omega}) >= 0$. This condition is satisfied by Bartlett's window but not by Hamming or Hanning windows; so although the later provide a better frequency resolution they cannot be used [16].

$$\varepsilon[S_{xx}(\omega)] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \varepsilon[I_N(\theta)] W(\exp^{j(\omega-\theta)} d\theta$$

Substituting for $\varepsilon[I_N(\theta)]$ we get

$$\varepsilon[S_{xx}(\omega)] = \frac{1}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} P_{xx}(\phi) W_B(\exp^{j(\theta-\phi)}) W(\exp^{j(\omega-\theta)}) d\phi d\theta$$

If M is small compared to N then $W(\exp^{j\omega})$ will be approximately constant compared to $W_B(\exp^{j\omega})$. Therefore

$$\varepsilon[S_{xx}(\omega)] = \frac{1}{2\pi}\int_{-\pi}^{\pi} P_{xx}(\theta)W(\exp^{j(\omega-\theta)})d\theta$$

We see that the windowing method is also a biased estimate. Increasing the spectral window width increases the smoothing but reduces frequency resolution of the spectrum. The covariance and variance of this estimate are [13]

$$Cov[S_{xx}(\omega_1), S_{xx}(\omega_2)] = \frac{1}{2\pi N}\int_{-\pi}^{\pi} P_{xx}(\phi)^2 W(\exp^{j(\omega_1-\phi)})W(\exp^{j(\omega_2-\phi)})d\phi$$

$$Var[S_{xx}(\omega)] = \frac{1}{2\pi N}\int_{-\pi}^{\pi} P_{xx}(\phi)^2 W(\exp^{j(\omega-\phi)})^2 d\phi$$

If W is narrow compared to variations in $P_{xx}(\omega)$ then

$$Var[S_{xx}(\omega)] = [\frac{P_{xx}(\omega)^2}{N}][\frac{1}{2\pi}\int_{-\pi}^{\pi} W(\exp^{j\phi})^2 d\phi] \tag{2.9}$$

We see that as N increases the variance approaches zero. Thus the windowing method also gives a consistent estimate of power.

Thus we see that these two smoothing techniques obtain a reduction in variance of the periodogram estimate. Also though both these techniques are biased estimators, they provide an asymptotically unbiased and consistent estimate of the spectrum. In the next chapter we wil see in detail the new proposed method for periodogram smoothing.

Chapter 3

The ATS method

3.1 Basic Theory

ATS technique is an approach to fitting curves and surfaces to data using a non-parametric regression procedure. The probability distribution of the data to be regressed may not necessarily be Gaussian. Suppose $z_i = (x_i, y_i)$ for $x_i = (x_{i1}, x_{i2}, ...x_{ir})$ for i=1 to n are measurements of r+1 variables. The x_i are r-dimensional data and y_i is some function of x_i. Let $y = g(x)$ be the functional relationship between x_i and y_i. For eg x_i may be a r-dimensional random sample and y_i its probability density function. y_i may be a binary response that depends on x_i.(In this case $g(x)$ may be $\varepsilon(y/x)$) For r=1 and x_i spectral frequencies, y_i may be the periodogram estimate of a time series. (This is the application which we are going to investigate) [8].

In the ATS method 'A' stands for averaging,'T' for transformation and 'S' for smoothing. This method can be employed only for those data distributions for which a suitable variance stabilization transformation('T') exists. For the first step a suitable value of 'a'(number of points to be taken for averaging) is chosen. The average $\bar{y}_i$ is calculated by taking 'a' values of y_k whose x_k are closest to x_i. The corresponding average $\bar{x}_i$ is also formed to get a set of data points of the form (x_i, y_i). To identify the x_k closest to x_i a suitable metric is to be chosen over the space of x_i. The averaging can be done either by using overlapping or non-overlapping type of windows. Next a variance stabilizing transformation 'T' is applied to the averaged

data to get another set of data $(\bar{x}_i, T(\bar{y}_i))$. In the final smoothing step 'S' the $(\bar{t}_i = T(\bar{y}_i))$ are smoothed as a function of $\bar{x}_i$ using a non-parametric regression procedure. If required an inverse transformation can be applied to undo the effect of the 'T' step and obtain an estimate of $g(x)$.

3.2 The Spectral Estimation problem

To define the problem consider a real discrete time random process $x[n]$. We consider a windowed sequence of length N. Then the periodogram estimate of power spectrum as defined section 2.1 is

$$I_N(\omega) = \frac{1}{N} \mid \sum_{n=0}^{N-1} x[n] \exp^{-j\omega n} \mid^2$$

Here we estimate the periodogram at discrete values of frequencies $\omega_i = \frac{i}{N}$ for i=0 to $[\frac{N}{2}]$. This is in fact the Discrete Fourier Transform(DFT). Let $y_i = I_N(\omega_i)$ be the estimates of the periodogram at frequencies $x_i = \omega_i$. If $g(\omega)$ is the spectral density to be estimated then we shall try and estimate the same using the ATS procedure described before. The power spectrum can be estimated for any value of frequency using this procedure. Here though we are assuming that the power spectrum is a smooth function of frequency.

3.2.1 Distribution of the estimate

We shall show that for $i = 1$ to $[\frac{N}{2}]$ the random variables $\frac{y_i}{g(\omega_i)}$ are chi-squared distributed with degree of freedom 2 and for ω_i=0 and $\omega_i = \pi$, $\frac{y_i}{g(\omega_i)}$ have degree of freedom 1 [13]. In case of real $x[n]$ we can rewrite the expression for periodogram estimate as

$$I_N(\omega) = \frac{1}{N}[(\sum_{n=0}^{N-1} x[n]cos(\omega n))^2 + (\sum_{n=0}^{N-1} sin(\omega n))^2]$$

We first consider the case when $x[n]$ is a white Gaussian noise sequence.

$$\begin{aligned} \varepsilon[x[n]] &= 0 \\ \varepsilon[x[n]^2] &= \sigma^2 \end{aligned}$$

$$\begin{aligned}
Let\, A(\omega) &= \sum_{n=0}^{N-1} x[n]cos(\omega n) \\
and\, B(\omega) &= \sum_{n=0}^{N-1} x[n]sin(\omega n) \\
Then\, \varepsilon[A(\omega)] = \varepsilon[B(\omega)] &= 0
\end{aligned}$$

For our problem since y_i are to be computed only at frequencies $\omega_i = \frac{2\pi i}{N}$ we consider only these frequencies. Therefore

$$\begin{aligned}
Var[A(\omega_i)] = \varepsilon[A(\omega_i)^2] &= \sigma^2 \sum_{n=0}^{N-1} cos^2(\frac{2\pi in}{N}) = \sigma^2 \frac{N}{2} \;\; k = 1\; to\; \frac{N}{2} - 1 \\
&= \sigma^2 N \;\; k = 0\; or\; \frac{N}{2} \\
Also\, Cov[A(\omega_k), A(\omega_l)] &= 0 \;\; k \neq l \\
And\, Cov[B(\omega_k), B(\omega_l)] &= 0 \;\; k \neq l \\
Cov[A(\omega_k), B(\omega_l)] &= 0 \;\; for\, all\, k\, and\, l
\end{aligned}$$

Since $A(\omega_i)$ and $B(\omega_i)$ are linear functions of normally distributed random variables they are also normally distributed. Therefore

$$\frac{A^2(\omega_i)}{Var[A(\omega_i)]} = \frac{2A^2(\omega_i)}{N\sigma^2} \;\; and \;\; \frac{B^2(\omega_i)}{Var[B(\omega_i)]} = \frac{2B^2(\omega_i)}{N\sigma^2}$$

are chi-squared distributed random variables with degree of freedom one (X_1^2). Therefore the sum $\frac{(A^2(\omega_i)+B^2(\omega_i))}{N/2\sigma^2}$ is chi-squared with degree of freedom two. For $i = 0$ or $\frac{N}{2}$ $B(\omega_i) = 0$ therefore $y_i = \frac{A^2(\omega_i)}{Var(A(\omega_i))}$ is chi-squared distributed with degree of freedom one. We now consider the case when $x[n]$ is not a white noise process. However a non-white process(coloured) can be obtained from a white one by suitable filtering operation. If $z[n]$ is a white noise process, $x[n]$ is the coloured noise process and $h[i]$ is the impulse response of the filter then

$$x[n] = \sum_{-\infty}^{\infty} z[i]h[n-i]$$

If $P_{xx}(\omega)$ and $P_{zz}(\omega)$ are the power spectrums of $x[n]$ and $z[n]$ respectively then

$$P_{xx}(\omega) = \mid H(\omega) \mid^2 P_{zz}(\omega) = \sigma_z^2 \mid H(\omega) \mid^2$$

Consider a finite record of length N of the coloured noise process.

$$I_N^x(\omega) = \frac{1}{N} \mid \sum_{n=0}^{N-1} x[n] \exp^{-j\omega n} \mid^2$$

$$= \frac{1}{N} \mid \sum_n \sum_k h[k] z[k-n] \exp^{-j\omega n} \mid^2$$

$$= \frac{1}{N} \mid \sum_n z[k-n] \exp^{j\omega(k-n)} \sum_k h[k] \exp^{-j\omega k} \mid^2$$

$$= \mid H(\omega) \mid^2 I_N^z(\omega)$$

$$\frac{I_N^x(\omega)}{\sigma_z^2 \mid H(\omega) \mid^2} = \frac{I_N^z(\omega)}{\sigma_z^2} \; ie \; \frac{I_N^x(\omega)}{P_{xx}(\omega)} = \frac{I_N^z(\omega)}{\sigma_z^2}$$

Since $\frac{I_N^z(\omega)}{\sigma_z^2}$ is X_2^2 for $\omega_i = \frac{i}{N}$ i=1 to $[\frac{N}{2}]$ and X_1^2 for ω_i=0 or$\frac{1}{2}$; therefore $\frac{I_N^x(\omega)}{P_{xx}(\omega)}$ also has the same distribution. Thus we have seen that the periodogram estimates obtained are independent of each other and are X_2^2 distributed. The frequencies other than $\omega = 0$ and $\omega = \frac{1}{2}$ having degrees of freedom 2 are exponentially distributed random variables. For the ATS operation we drop the end frequencies 0 and $\frac{1}{2}$ and consider only the other frequencies.

3.3 The averaging step 'A'

In the averaging step we average the periodogram in 'b' blocks of 'a' consecutive Fourier frequencies. We use averaging with disjoint windows and since all quanties are unidimensional the metric used is arithmetic difference of the quantities under consideration. The average $\bar{y}_l$ for the l^{th} block is associated with the averaged frequency $\bar{x}_l$. The number of blocks thus obtained is $b = [\frac{((n-1)/2)}{a}]$. Since all the quantities averaged are exponential random variables their average will be gamma distributed. Averaging with disjoint windows maintains the independence of the estimates.

3.3.1 Bias and variance of the averaged estimates

The averaging done in the 'A' step is equivalent to the windowing technique of periodogram smoothing. Here a rectangular window(in the frequency domain) is used.

This leads to reduction in variance of the estimate, the reduction being proprtional to the value of 'a'. Thus the expressions of bias and variance derived for the windowing step are valid for this. However since we are computing the periodogram only at discrete values of frequency the integration operation is replaced by the summation operation. Also the bias and variance are proportional to the frequency ω_i. The expressions for mean and variance as functions of the frequency ω_i are [15]

$$mean(\omega_i) = \frac{P_{xx}(\omega_i)}{2\pi} \sum_{\theta=-\frac{(a-1)}{2}}^{\frac{(a-1)}{2}} W(\exp^{j\theta}) \tag{3.1}$$

$$Var(\omega_i) = \frac{P_{xx}^2(\omega_i)}{N}[\frac{1}{2\pi} \sum_{\theta=-\frac{(a-1)}{2}}^{\frac{(a-1)}{2}} W^2(\exp^{j\theta})] \tag{3.2}$$

Here $W(\exp^{j\theta})) = \frac{1}{a}$ for all θ. From the above expressions we see that the mean and variance depend on the frequency ω_i. Also the variance is related to the mean by the relation $Var(\omega_i) = Kmean^2(\omega_i)$ for a suitable constant K. Since the mean and variance depend on the power content at that frequency, for a non-white process the variance will change with the frequency. Regression analysis is used to fit a curve of the form $y_i = g(x_i) + \varepsilon_i$ where ε_i are i.i.d random variables. Thus it is necessary to investigate whether there exists a transformation of some kind which will make the variance of the estimate a constant at all frequencies which is independent of the frequency and the power content at that frequency.

3.4 The variance stabilizing transform 'S'

The purpose of this transformation is to change the scale of the observed data to make the data analysis more valid. The transformed variate should also have a constant variance. Ideally the transformed variable must have the following properties

- Variance of the transformed variable must be unaffected by changes in the mean level.
- The transformed variable should be normally distributed.

- The transformed scale should be one in which the arithmetic average is an efficient estimate of the mean.

- The real effects should be linear and additive.

The required transformation function depends on the mean, variance and distribution of the random variable. Consider a random variable x whose mean $\varepsilon(x)$ is a real variable having the range R. If σ_x^2 is the variance of x, let it be related to the mean m through the relation $\sigma_x = f(m)$. Thus any fluctuations in the mean are going to affect the variance. In case of averaged periodogram the mean being a function of the power at that frequency, the variance is also then a function of power. In order to stabilize the variance ie make it a constant irrespective of the power or frequency it is necessary to transform the variable x through a transformation $t(x)$ such that the variance of the transformed variable is independent of the mean. This means that we should achieve

$$\frac{\partial y}{\partial m} = 0 \; and \; \frac{\partial \sigma_y^2}{\partial m} = 0$$

for all m in R. Also starting from $\frac{dt}{dx} = t'(x)$ it can be deduced by a summation procedure that $\sigma_y = t'(m)\sigma(m)$. This derivation is not mathematically sound and can be used if on application it gives satisfactory results. For our case we shall use this result to derive the necessary transformation function.

After application of 'A' step the relation between mean and variance is

$$\sigma(m) = Km$$

where K is a constant. If $y = t(x)$ is the transformation being used then it is required that $Var(y) = \sigma_y$=constant(say c). Thus we have

$$\sigma_y = (\frac{dt}{dx})_{x=m} Km$$

ie

$$c = \frac{dt}{dm} Km$$

$$dt = K' \frac{dm}{m} \; or \; t = K' log(m)$$

In general we have $t(x) = Klog(x)$. Thus for a chi-squared periodogram estimate which is smoothed by a window function the variance stabilizing transform is the logarithm function. More details on the logarithm transformation can be found in [9] [3]. As stated before the transformation function depends on the distribution of the random variable. In sec 3.1 we had stated that one of the other applications of this method is the case when y is a binary response on x. Here the required transformation is the inverse sine function [12].

3.5 The smoothing process 'S'

The final step consists of smoothing by regression. We have the transformed variables $t_i = log(\bar{y}_i)$ which are i.i.d random variables. We use a regression technique known as locally weighted regression(loess) [6].

Locally weighted regression or loess is a method of estimating a regression surface through a multivariate smoothing procedure. It consists of fitting a function of independent variables locally and in a moving fashion analogous to a moving average filter. Loess is different from the usual polynomial regression. Loess can be used to estimate a wider class of regression surfaces than the polynomial method. Since the spectum smoothing is an univariate problem we shall discuss only the univariate loess method. The same technique has been generalized and used for the multivariate case [7].

Consider a set of points of the form (x_i, y_i). By applying locally weighted regression we shall estimate a new set of points $(x_i, \hat{y}_i)$. This new set of points $\hat{y}_i$ are the fitted values at x_i. If the regression surface to be estimated is of the form $y_i = g(x_i) + \varepsilon_i$ then $\hat{y}_i = g(x_i)$.

In order to describe the smoothing procedure we consider a window function having the following properties

1. $W(x) > 0 \mid x \mid < 1$
2. $W(-x) = W(x)$
3. $W(x)$ is a non-increasing function for $x >= 0$

4. $W(x) = 0 \mid x \mid > \cancel{0}\ 1$

Let f be a parameter($0 < f <= 1$) used to control the amount of smoothing in the regression process. If U is the number of points of the form (x_i, y_i) which are to be smoothed; let $r = Uf$ be rounded off to the nearest integer. For each x_i the weights $w_k(x_i)$ are defined for k=1 to U. This is done by centering the weight function W at x_i and scaling it so that the point at which it first becomes zero is the r^{th} nearest neighbour of x_i. Then y_i is determined as the value of a d^{th} degree polynomial fit to the data using weighted least squares with weights $w_k(x_i)$. This procedure is carried out for every x_i to obtain the values $\hat{y}_i$. After this a set of residuals $\delta_i = y_i - \hat{y}_i$ is defined for each x_i. New fitted values are computed using the same procedure just described but using a new set of weights $\delta_i w_k(x_i)$. This procedure is carried out for a specified number of iterations after which a smooth fit is obtained.

The assumption of smoothness allows us to use the points in the neighbourhood of (x_i, y_i) to determine $\hat{y}_i$. The values of the weight function $w_k(x_i)$ decreases as the distance from x_i increases. Thus the points which are close to x_i play a larger role in determining the value of $\hat{y}_i$ than those points which are farther away from x_i. Increasing the value of 'f' increases the neighbourhood of influential points and therefore increases the smoothness of the fit.

We shall now breifly describe the algorithm used in the loess procedure. For each x_i, h_i is the distance of the r^{th} nearest neighbour of x_i ie h_i is the r^{th} smallest value of $\mid x_i - x_k \mid$ for k=1 to $\underset{U}{\cancel{S}}$. The weighting function is defined as

$$w_k(x_i) = W(\frac{x_k - x_i}{h_i})$$

The following sequence of operations is carried out

1. At the point (x_i, y_i) we are fitting a d^{th} order polynomial to obtain the value of $\hat{y}_i$ as $\hat{y}_i = \sum_{j=0}^{d} \beta_j(x_i) x_i^j$. The estimates of the polynomial coefficients $\hat{\beta_j}(x_i)$ are determined using the method of weighted least squares. The βs' are obtained by minimizing the value of

$$\sum_{k=1}^{\cancel{S}\,U} w_k(x_i)(y_k - \beta_0 - \beta_1 x_k \ldots \beta_d x_k^d)$$

Then

$$\hat{y}_i = \sum_{j=0}^{d} \hat{\beta_j(x_i)} x_i^j$$

This is carried out for all x_i.

2. Let B be a bisquare weight function which is defined by

$$\begin{aligned} B(x) &= (1-x^2)^2 \mid x \mid < 1 \\ &= 0 \mid x \mid >= 1 \end{aligned}$$

If $\delta_i = y_i - \hat{y}_i$ are the residuals from the fitted values; let 'p'be the median of these fitted values $\mid e_i \mid$. We define robust weights as $\delta_k = B(\frac{e_k}{6p})$.

3. The steps (1) and (2) are repeated using the weights $\delta_k w_k(x_i)$. Thus a new set of smoothed points y_i is obtained. This procedure is carried out for a certain number of iterations.

The weight function is derived from the tricube function

$$\begin{aligned} W(x) &= (1- \mid x \mid^3)^3 \mid x \mid < 1 \\ &= 0 \mid x \mid >= 1 \end{aligned}$$

The above procedure can be used to determine the value of $\hat{y}$ at any value of x(ie other than x_i). Thus a smooth and continuous curve of the power spectrum can be obtained.

3.6 Choice of various parameters in the loess algorithm

There are four parameters which are to be considered in the algorithm. They are

1. The order of the polynomial fit,'d'.
2. The weighting function to be used.'w'.
3. The number of iterations,'t'.

4. The smoothing parameter,'f'.

The choice of 'd' is a trade-off between the need to reproduce variations in the spectrum and computational ease. The simulations done by us are performed for $d = 2$. However the algorithm can be used for any value of 'd'.

The requirements of the weighting function 'w' have been stated before. The reasons are stated now. The first requirement is chosen because negative weights do not make any sense. The second requirement $W(-x) = W(x)$ is chosen to give the same importance to points which are at the same distance on both sides of x_i. $W(x)$ should be non-increasing with increasing x so that a point nearer to x_i will have more weight than a point which is farther from x_i. In addition $W(x)$ should decrease smoothly from $x = 0$ to $x = 1$. The tricube function is chosen because it enhances a chi-squared distributional approximation of the estimate of error variance [6]. In our simulations we have used two iterations. It has been studied that two iterations give a good, robust fit. However some type of convergence criterion can be defined for the loess algorithm. Execution of the algorithm is carried out till this criteion is met. This method however has not been formulated or implemented.

It had been stated earlier that this method can be used only for cases where the spectrum is a smooth function frequency. However the extent of smoothness of the curve is not known and no procedure has been adopted to measure it. Broadly it can be stated that the amount of smoothing must be enough to remove the fluctuations in the spectrum due to the statistical nature of the data and finite record length. But it should not be excessive so as to distort or smoothen out the actual variations in the spectrum. For eg for a narrow band signal whose power is concentrated in a narrow band of frequencies smoothing may be dangerous for the spectrum. The simulations were therefore carried out for different smoothing factors and the plots compared.

3.7 Scaled Exponential Variables

To illustrate the problem of smoothing a periodogram consider the problem of smoothing scaled exponential variables [8]. Let

$$y_i = \mu_i W_i \quad i = \frac{-(n-1)}{2}\ to\ \frac{(n-1)}{2}$$

where W_i are independent standard exponentials and μ_i are smooth functions of i. Let $\mu_i = \exp(\alpha + i\beta)$

We assume that the objective is to estimate the centre value of μ ie $\exp(\alpha)$. If we assume that the record length n is of the form $n = ab$ then we can divide the data in 'b' disjoint blocks and apply the 'A' step that averages the 'a' values within each block. In the 'T' step we take the logarithms and in the 'S' step we shall perform smoothing by averaging by averaging these 'b' quantities. This process should give a good estimate of α. Since $\exp(\alpha + i\beta)$ is symmetric about the centre value, smoothing by averaging gives the same estimate as obtained by a linear fit. After the 'A' and 'T' steps we get

$$y_j = \alpha + \beta a j + ln(\hat{f}) + ln(\sum_{k=\frac{-(a-1)}{2}}^{(a-1)} 2W_{jk}/a) + Z_j$$

where $W_{jk} = W_{j+k}, f_k = \exp^{k\beta}, \hat{f} = \sum_{k=\frac{-(a-1)}{2}}^{\frac{(a-1)}{2}} \exp^{k\beta}, g_k = \frac{f_k - \hat{f}}{\hat{f}}, V_{jk} = \frac{W_{jk}}{\sum_k W_{jk}}, Z_j = ln(1 + \sum_k g_k V_{jk})$.

The value of j varies from 1 to 'b'. The 'S' step gives $\hat{y} = ave(y_j)$. We shall now derive the mean and variance of $\hat{y}$ and y_j. Since W_i are independent exponentials $\sum_{k=\frac{-(a-1)}{2}}^{\frac{(a-1)}{2}} W$ is gamma(a) distributed. Then [1]

$$\varepsilon[ln(\sum_k W)] = \psi(a)$$

$$Var[ln(\sum_k W)] = \psi'(a)$$

$\psi(a) = \frac{gamma'(a)}{gamma(a)}$, $\psi'(a)$ is the derivative of $\psi(a)$ and gamma'(a) is the derivative of gamma(a). To calculate the moments of Z_j we expand it by its Taylor's series expansion.

$$ln(1 + \sum_k g_k V_{jk}) = \sum_{r=1}^{\infty} \frac{(-1)^{(r-1)}(\sum_k gV)^r}{r}$$

Using the independence of V and $\sum W$ we have

$$\varepsilon(\sum_k gV)^r = \frac{\varepsilon(\sum gW)^r}{\varepsilon(\sum W)^r}$$

Consider the expression

$$\sum_{r=0}^{\infty} \frac{\theta^r}{r!}\varepsilon(\sum_k gW)^r = \varepsilon(\sum_{r=0}^{\infty} \frac{\theta^r(\sum_k gW)^r}{r!})$$

$$= \varepsilon(\exp(\theta \sum_k gW)) = \Pi_k(1-\theta g)^{-1} = \sum_r \theta^r h_r$$

Here h_r is the r^{th} homogeneous symmetric polynomial in the $g's$ which can be expressed in terms of the power sums $s_r = \sum_r g^r$ [11]. Also we have from the above expressions

$$\varepsilon(\sum_k gW)^r = r!h_r$$

From the tables in [11] we get $h_1 = s_1 = 0, 2h_2 = s_2, 6h_3 = 2s_3, 24h_4 = 6s_4 + 3s_2^2$. Thus we get

$$\varepsilon(Z) = \frac{-s_2}{2a^{[2]}} + \frac{2s_3}{3a^{[3]}} - \frac{3(2s_4+s_2^2)}{a^{[4]}} + \ldots$$

$$\varepsilon(Z^2) = \frac{s_2}{a^{[2]}} - \frac{2s_3}{a^{[3]}} + \frac{11(2s_4+s_2^2)}{4a^{[4]}} + \ldots$$

$$a^{[r]} = a(a+1)(a+2)\ldots(a+r-1)$$

From these we get the moments of y_j and $\hat{y} = ave(y_j)$ as

$$\varepsilon(\hat{y}) = \alpha + ln(\hat{f}) + \psi(a) - ln(a) + \varepsilon(Z)$$

$$nVar(\hat{y}) = a\psi'(a) + aVar(Z)$$

It turns out that for small values of 'a' (a ≤ 8 approx) the contribution of $\varepsilon(Z)$ and $Var(Z)$ to the mean and variance of $\hat{y}$ is negligible. Thus the mean and variance of $\hat{y}$ mainly depend on $\psi(a)$, $\psi'(a)$ and 'a' itself (for small values of 'a'). Also the variance of t_j ie the y_j is approximately

$$Var(t_j) = \psi'(a)$$

Thus we see that the variance of the transformed variables t_j is a constant which is independent of their distributions. These stabilized variances have been verified through simulations.

Chapter 4

Simulation Results and Conclusions

This chapter contains the simulations carried out and the conclusions inferred from these simulations.

4.1 Generation of the Random Signal

The first task is the generation of the necessary random signal (data). For this the pseudo random number generator algorithm of Park and Miller was used [18]. This algorithm generates a pseudo random number which is uniformly distributed between 1 and $2^{31}-2$. This number is then divided by $2^{31}-1$ to generate a random number which is uniformly distributed between 0 and 1.(exclusive of the end-points) This uniformly distributed number can then be used to generate a wide range of distributions. For this purpose we make use of the following theorem

If U is uniformly distributed in [0,1] then $F^{-}(U)$ is a random variable whose cumulative distribution function is F.

Using this we can generate different distributions provided their cumulative distribution function is known and is invertible. Thus in order to generate an exponential random variable whose probability distribution function is $f(x) = \lambda\exp(-\lambda x)$ the necessary function $F^{-}(U)$ is $\frac{-1}{\lambda}ln(1-U)$. However since U is uniformly dis-

tributed so is $(1 - U)$. Therefore the required exponential random variable can be generated using $\frac{-1}{\lambda} ln(U)$. The data to be generated has to be initialised with a seed value which is a negative integer. The average of the entire record gives an estimate of the mean which can be substracted from each value to get a zero mean random variable. For a zero mean random variable the autocovariance is the same as autocorrelation. For different seeds the data stream starts from different numbers. Similarly other types of data can also be generated.

4.2 Simulations

4.2.1 Simulations for the Stabilizing Transform

In sec 3.4 we had learnt about the variance stabilizing transformation. We had seen that the ATS method can only be used if such a transformation exists. Also we had seen that for the periodogram the required transformation is the log function. We shall verify through simulations if this transformation results in a constant variance that it claims. For this purpose it is necessary to perform the simulations independently a large number of times and the variance must be calculated from these simulations. Here we are assuming that the time averages will give us a good estimate of the variance. The following experiment was carried out

1. For i=1 to M do upto step 6
2. Using seed value as -i generate uniform random numbers $x[n]$ in [0,1] for n=1 to N.
3. Generate $y[n] = \sum_{j=1}^{K} a_k y[n-k] + x[n]$ for $n = 1$ to N.
4. Calculate the FFT of $y[n]$ at frequencies $\omega_k = \frac{k}{N}$ for $k = 0$ to $\frac{N}{2}$ and estimate the power spectrum from this using eq.2.3.
5. Divide the entire frequency range into 'b' blocks of 'a' frequencies each. $b = \frac{(N/2-1)}{2}$. Find the average value of the estimate and the average value of the frequency. Let I_N denote the estimate, then store this result

in $I_N[j,i]$ which indicates that this is the power spectrum estimate for the j^{th} frequency for the i^{th} seed value. j varies from 1 to b.

6. Take the logarithm of $I_N[j,i]$ and store it in $T_N[j,i]$.

Then the variance before the transformation at frequency j is calculated as

$$\sigma^2_{I,j} = \frac{1}{M}\sum_{i=1}^{M}(I_N[j,i] - \frac{1}{M}\sum_{i=1}^{M} I_N[j,i])^2$$

Similarly the variance after the log transform $\sigma_{T,j}$ can also be calculated by using $T_N[j,i]$ in place of $I_N[j,i]$ in the above formula. The above simulations are carried out for different values of averaging 'a'. In our simulations the record lenth N was taken to be 1024. The variance calculations were done by averaging over 10000 independent iterations. The data $y[n]$ was generated from a uniformly distributed data $x[n]$ using a second order AR process with $a_1 = 0.0$ and $a_2 = 0.9025$. This data has power spectrum as shown in fig.4.1.

Fig.4.2 is the plot of variance versus frequency for averaging value a=2. Fig.4.2(a) shows the variance before the transform while fig.4.2(b) shows the figure after the transform. We can see that variance before the transform has the same profile as that of the power spectrum(see fig.4.1) while the variance after the transform is approximately constant over the entire frequency range. This constant variance is equal to 0.64 which agrees with the theoritical value

$$\psi'(a) = 0.6445 \;\; at \;\; a = 2$$

Fig.4.3 shows the plots for a=4. Here the variance after the transform stabilizes at 0.28. Theoritically we have

$$\psi'(a) = 0.28375 \;\; at \;\; a = 4$$

4.2.2 Simulations for Periodogram Smoothing

The next simulations are for the smoothing of the periodogram estimate. The smoothing was done using the ATS and the Bartlett's method. Here too the operations carried out are the same till the log transform step. After that the loess

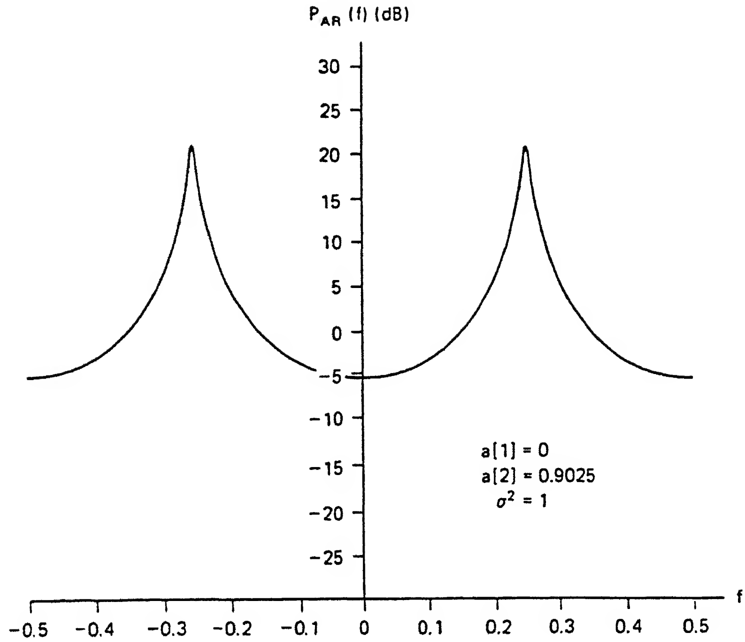


Figure 4.1: Power Spectrum of the AR2 Process y[n]

algorithm is used for smoothing the spectrum. The variables in the loess algorithm are the smoothing factor 'f' ,degree of the polynomial to be fit 'd' and the number of iterations 't'. In order to estimate the spectrum by Bartlett's method we divide the record length N in 'a' disjoint records so as to have $\frac{N}{a}$ as the length of each. The periodogram estimate of each block is estimated independently and the average of the 'a' periodograms is the Bartlett's estimate. Here 'a' is the averaging factor used in ATS. We have used a record length N of 4096 data points. The degree of the polynomial fitted is 2 and the number of iterations carried out also is 2.

Figs.4.4 to 4.6 are the plots for estimating the spectrum of exponential white noise.(with mean arrival rate $\lambda = 1$) Fig.4.4 is for 'a'=2. In figs.4.4 (a), (b) and (c) the smoothing factors used are 0.01, 0.05 and 0.1 respectively. Fig.4.4 (d) is the Bartlett's estimate as the average of two disjoint periodograms. Figs.4.5 and 4.6 are the plots of the simulations carried out for 'a'=4 and 'a'=8 respectively. Next the simulations were carried out on a narrow-band signal whose power spectrum is as shown in fig.4.1. The data was generated using the iterative procedure

$$y[n] = -0.9025y[n-2] + x[n]$$

where $x[n]$ is uniform in [0,1]. Figs.4.7 and 4.8 are the plots for a=2. Fig.4.7 are the plots of the estimate using the ATS method for four smoothing factors of 0.01, 0.05, 0.1 and 0.3. The fourth smoothing factor has been included to show how excess

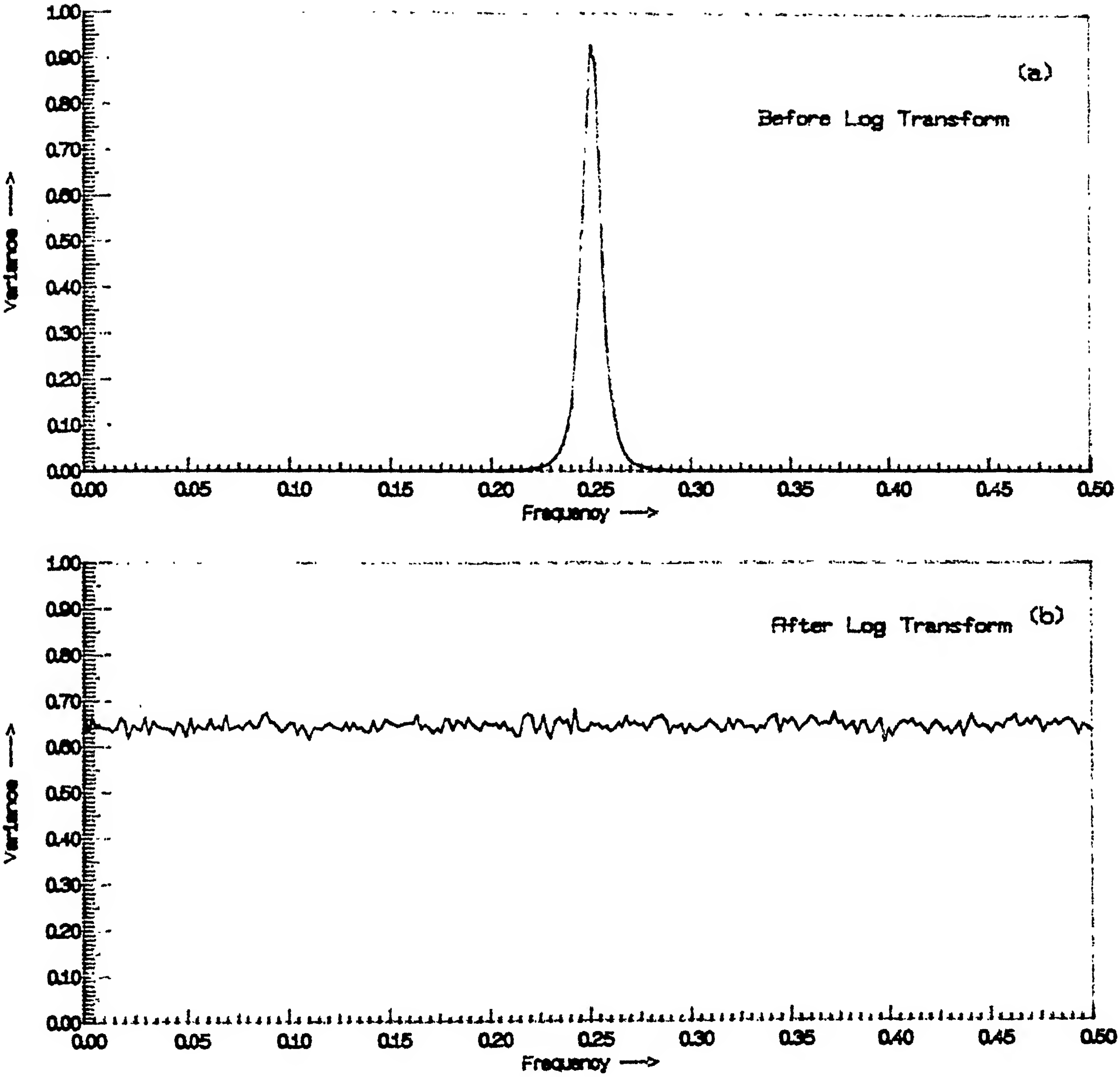


Figure 4.2: Estimate Variance vs Frequency.(a=2)

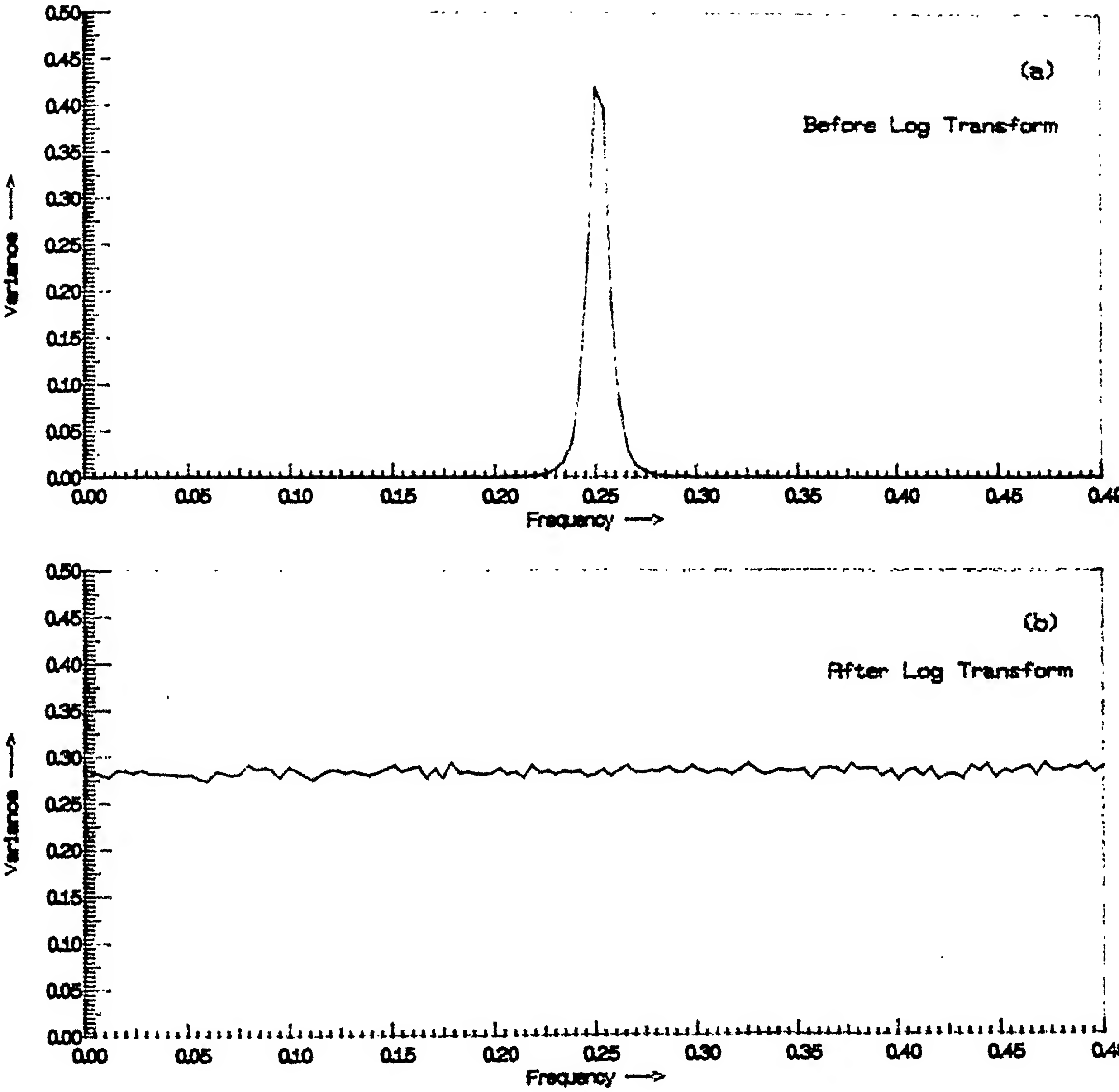


Figure 4.3: Estimate Variance vs Frequency.(a=4)

smoothing can distort the spectrum. Fig.4.8(a) is the plot of the spectrum after the 'A' and 'T' steps so as to compare it with the estimates obtained after different values of smoothing. Fig.4.8(b) is the Bartlett's estimate. Figs.4.9 and 4.10 are the plots for a=4 while figs.4.11 and 4.12 are for a=8. The first case of estimating the spectrum of white noise was a smooth spectrum while the second spectrum had some variations in it. In the third case we shall take a spectrum having sharp peaks or discontinuities in it. This gives us a fairly good idea of how the method works for a smooth spectrum at one end to a spectrum having sharp peaks at the other end. Therefore the third case we estimate the spectrum of sinusoids buried in white noise.

$$x[n] = sin(\omega_1 n) + sin(\omega_2 n) + v[n]$$

where $v[n]$ is uniform in [-0.5.0.5].

$$\omega_1 = 0.5 \, and \, \omega_2 = 0.6$$

In the plots the x-axis has the normalized frequency. Therefore $\omega_1 = 0.5$ and $\omega_2 = 0.6$ correspond 0.0796 and 0.0955 on the X-axis respectively. Fig.4.13 shows the plots using the ATS method using smoothing factors 0.01, 0.05 and 0.1 respectively. Here we see that the smoothing is not able to resolve the two sinusoids. In fact fig.4.14(a) which has the spectrum after the 'A' and 'T' steps the two sinusoids were clearly resolved. Smoothing thus has distorted the spectrum. The Bartlett's method fig.4.14(b) gives a good estimate. The method can also be tried on Gaussian distributions.

4.3 Conclusions

We had seen before that the ATS method can only be used in cases where a variance stabilization transformation exists. The importance of this transform has been verified through simulations. Due to the transform the estimates have the same variance at all frequencies and hence a regression procedure can be used. Regression is used to fit a model of the type $y_i = g(x_i) + \varepsilon_i$ where ε_i are i.i.d. and have a constant variance. The log transform effects this constant variance.

As seen from the various plots the ATS method is quite efficient when the power spectrum is a smooth function of frequency. As seen in the case of the exponential white noise this method gives a better estimate than the Bartlett's method. The Bartlett's method estimate has a lot of fluctuations due to the statistical nature of the data and due to the finite record length. Regression helps to smooth these fluctuations. In the case of the narrowband spectrum the ATS method gives better results for small values of the smoothing factor 'f'. At these values the estimate is better than the Bartlett's estimate. However for large values of smoothing factor it distorts the spectrum. Compared to this the Bartlett method gives better results in the sense that it does not distort the actual variations in the spectrum. However if the spectrum has sharp peaks then Bartlett's method can be preferred over the ATS method.

Thus we can conclude that the ATS method gives good results(better than Bartlett's method) if the spectrum to be estimated is a smooth function of frequency.

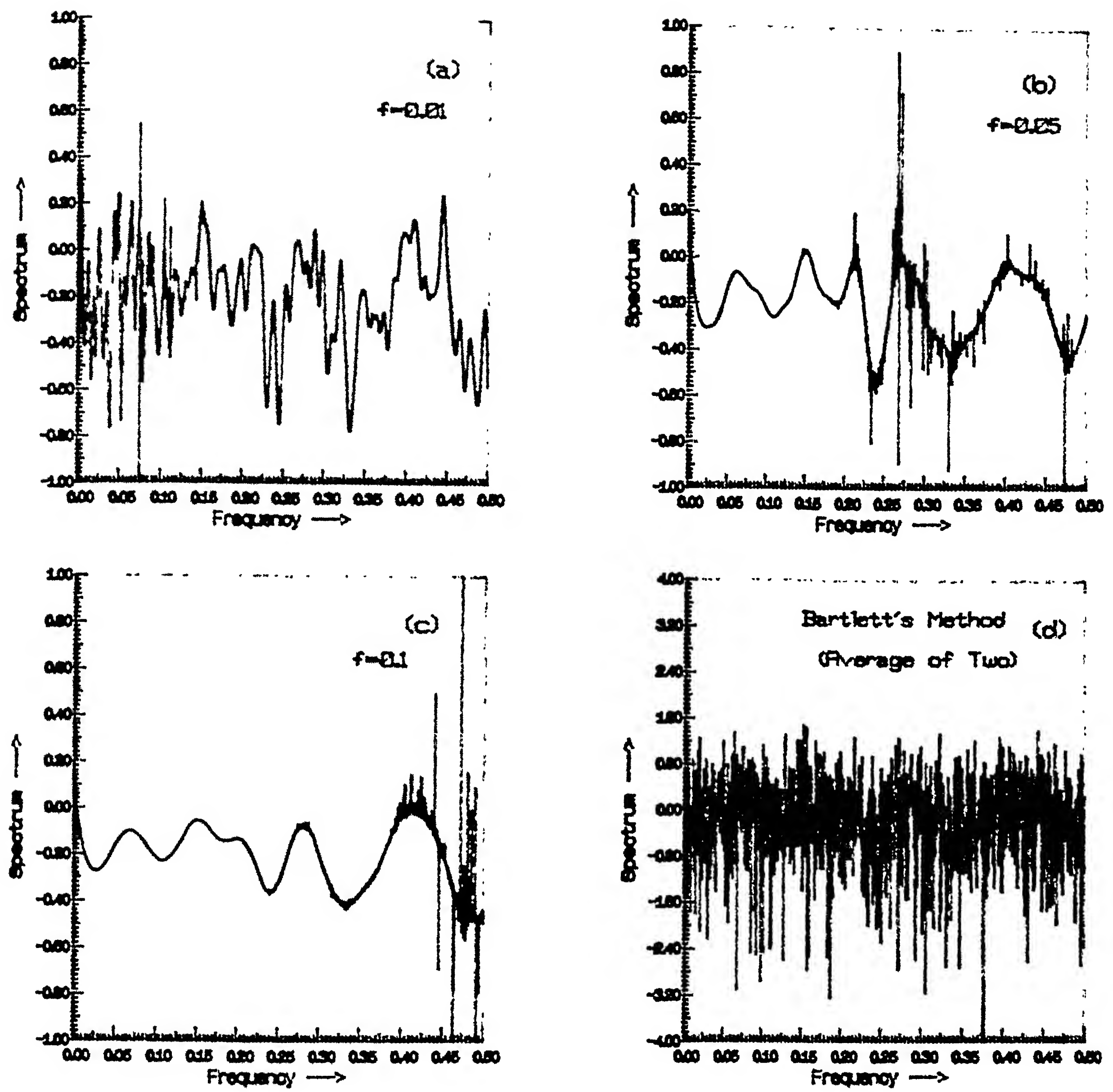


Figure 4.4: Exponential White Noise Spectrum Estimate by ATS and Bartlett's Methods.(a=2)

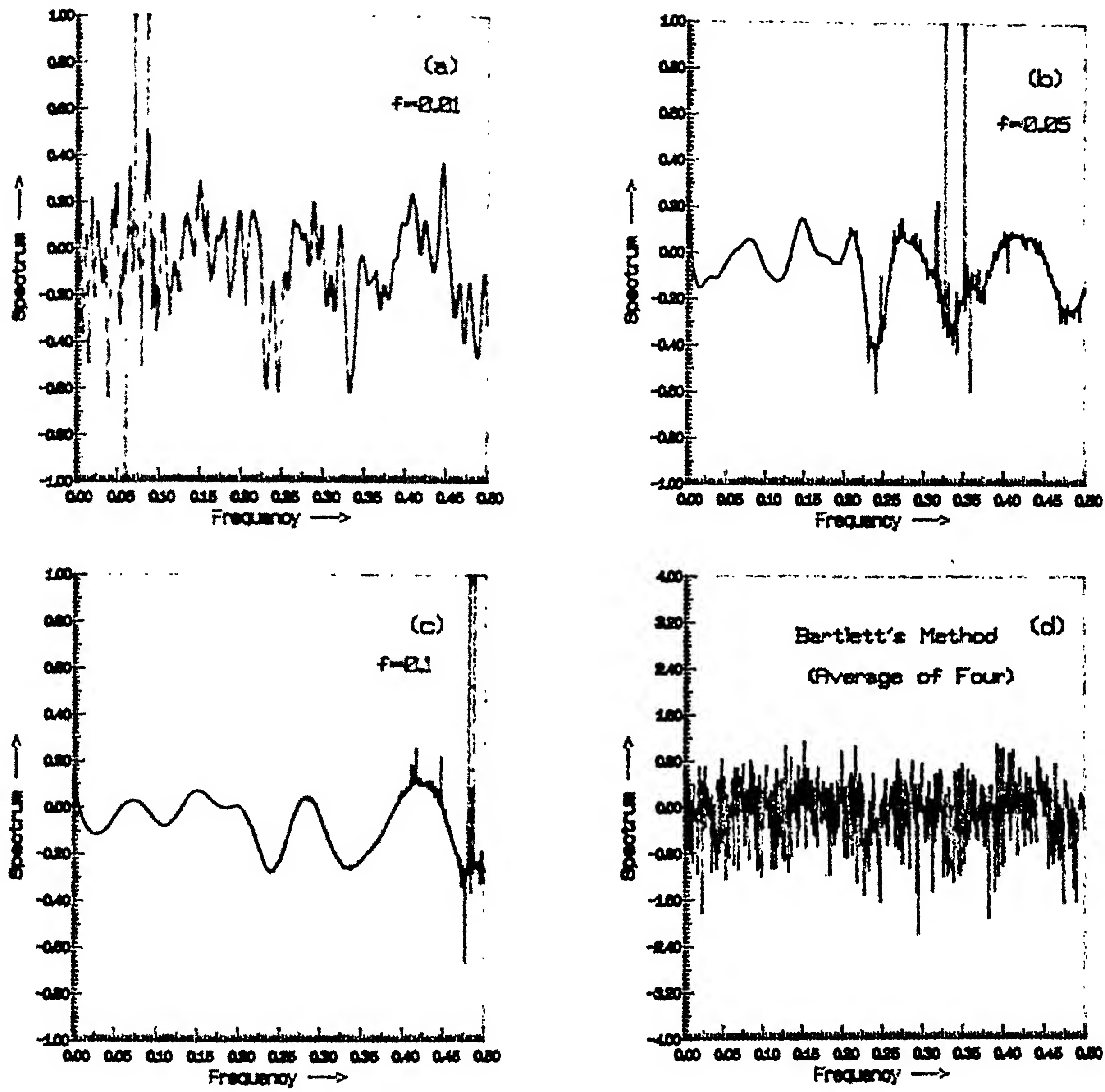


Figure 4.5: Exponential White Noise Spectrum Estimate by ATS and Bartlett's Method.(a=4)

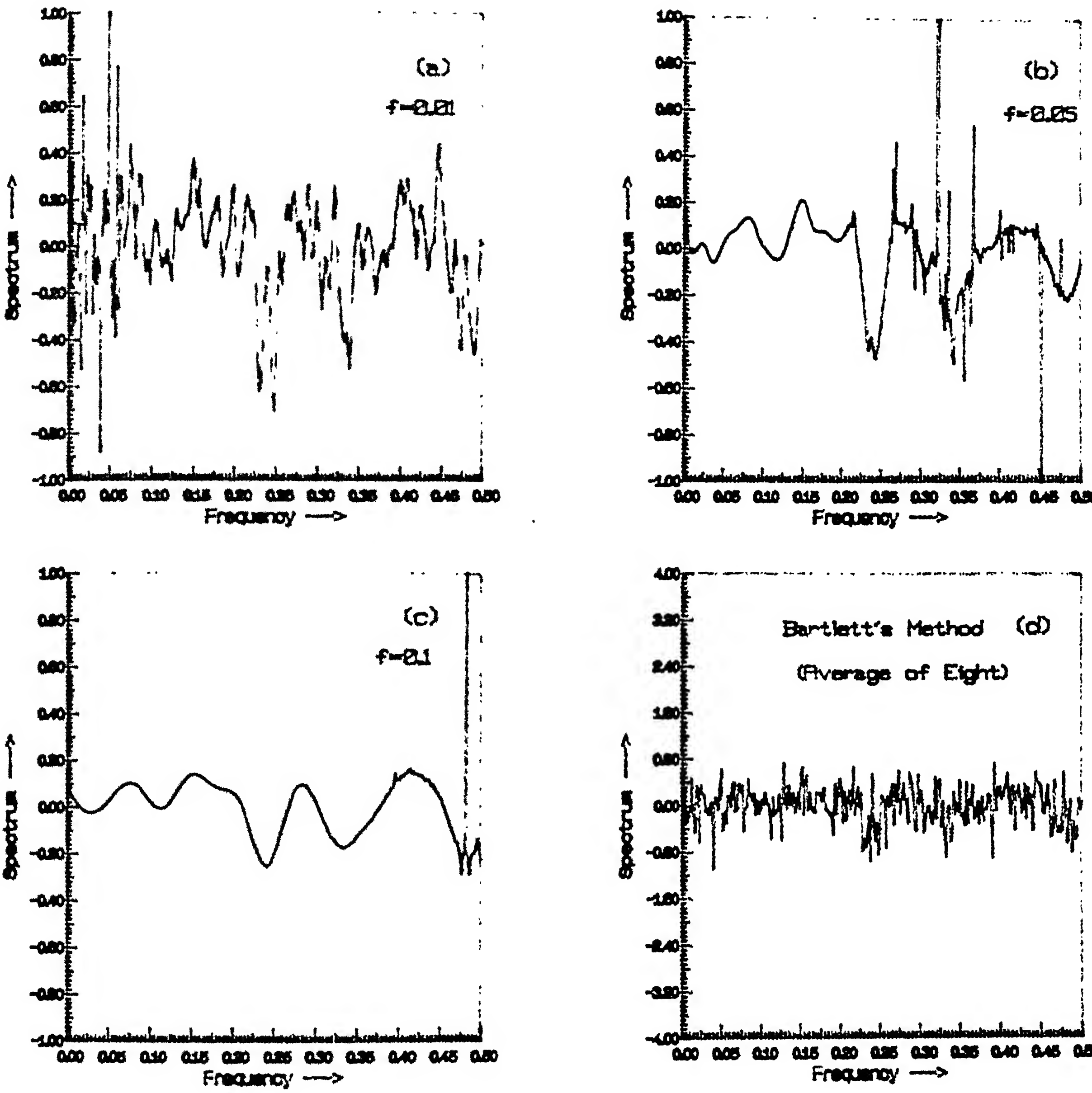


Figure 4.6: Exponential White Noise Spectrum Estimate by ATS and Bartlett's Method.(a=8)

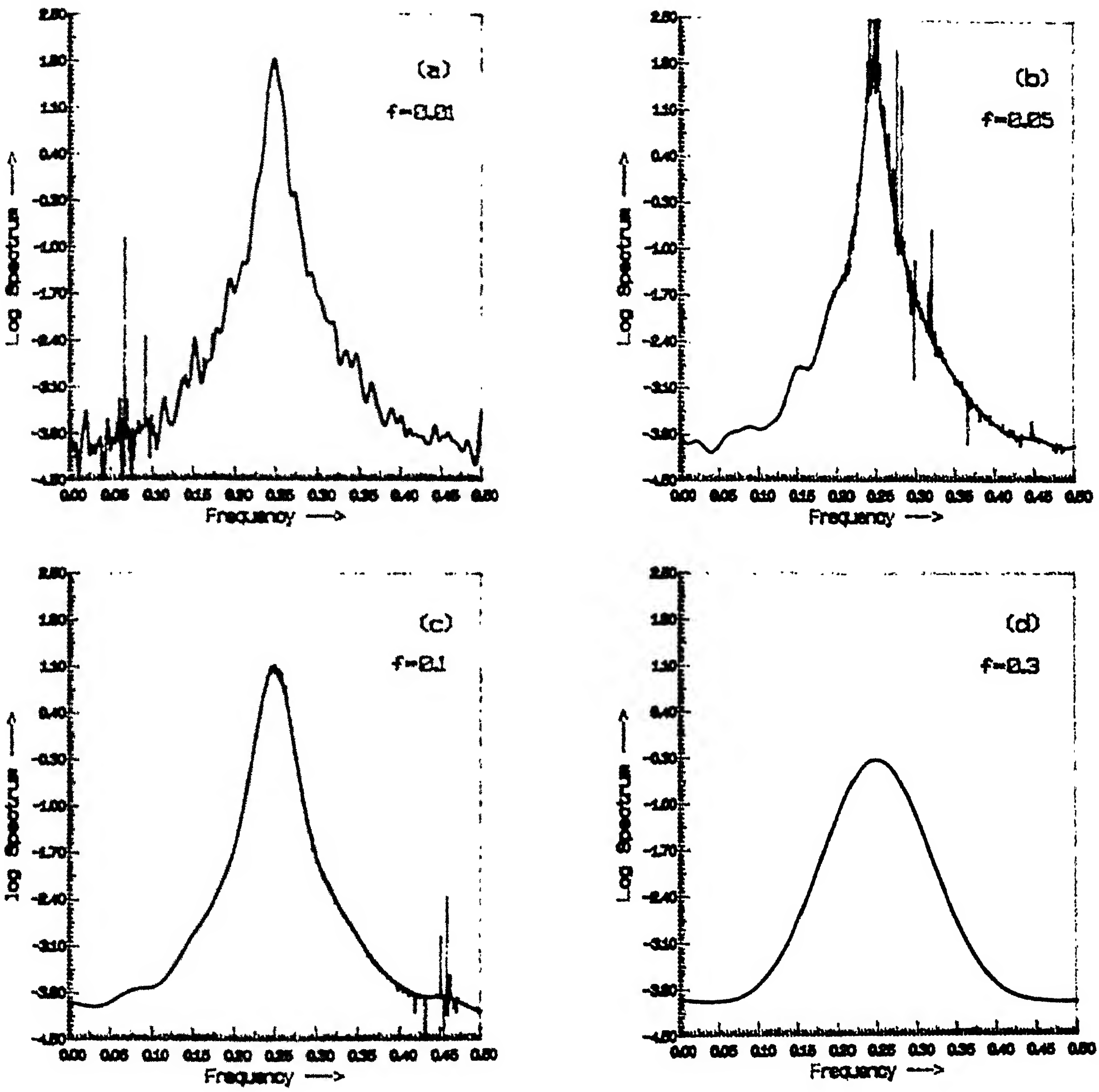


Figure 4.7: Spectrum Estimate of a Narrow Band Signal by ATS Method.(a=2)

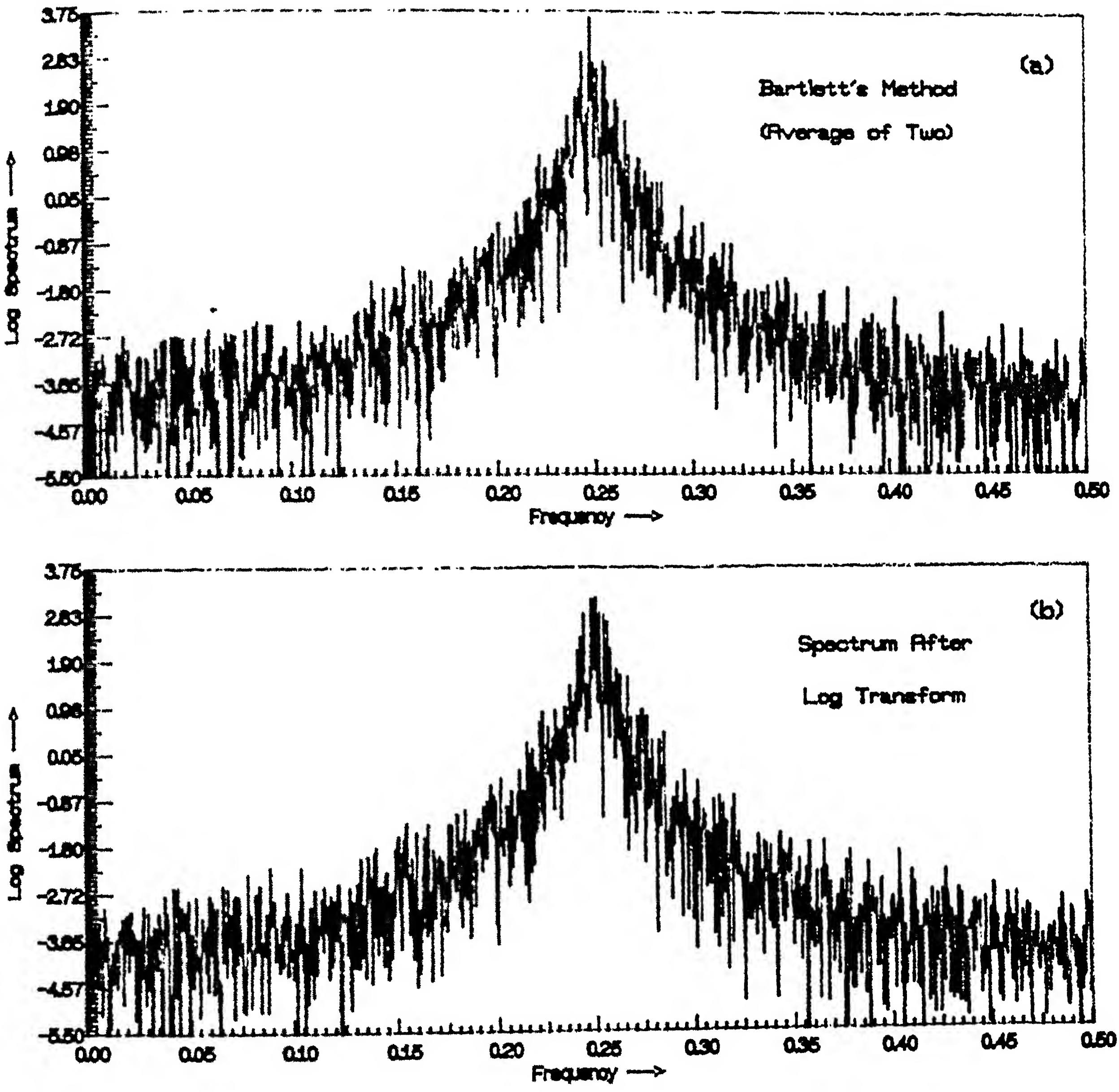


Figure 4.8: Spectrum Estimate of a Narrow Band Signal by Bartlett's Method. (Average of Two)

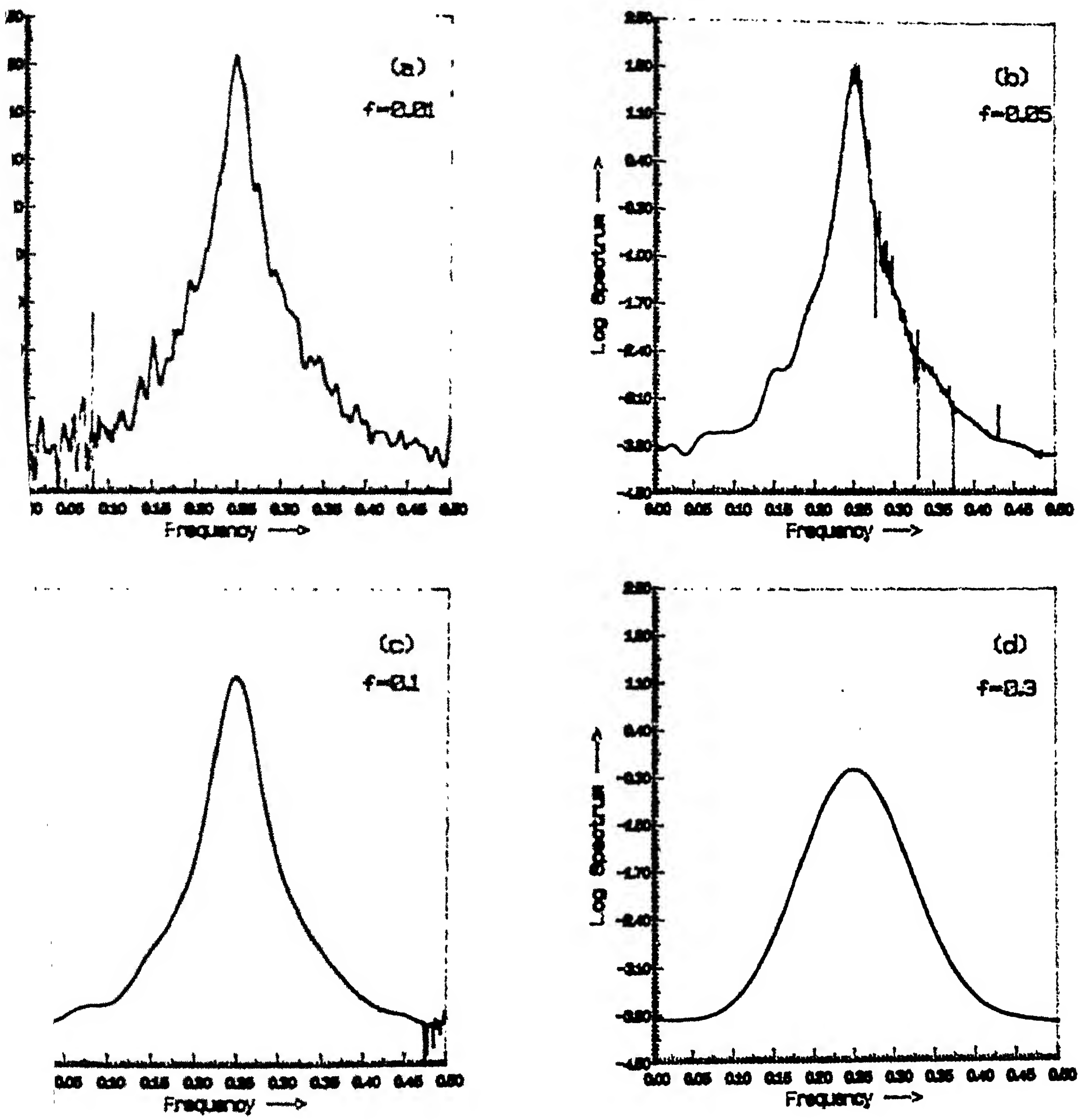


Figure 4.9: Spectrum Estimate of a Narrow Band Signal by ATS Method.(a=4)

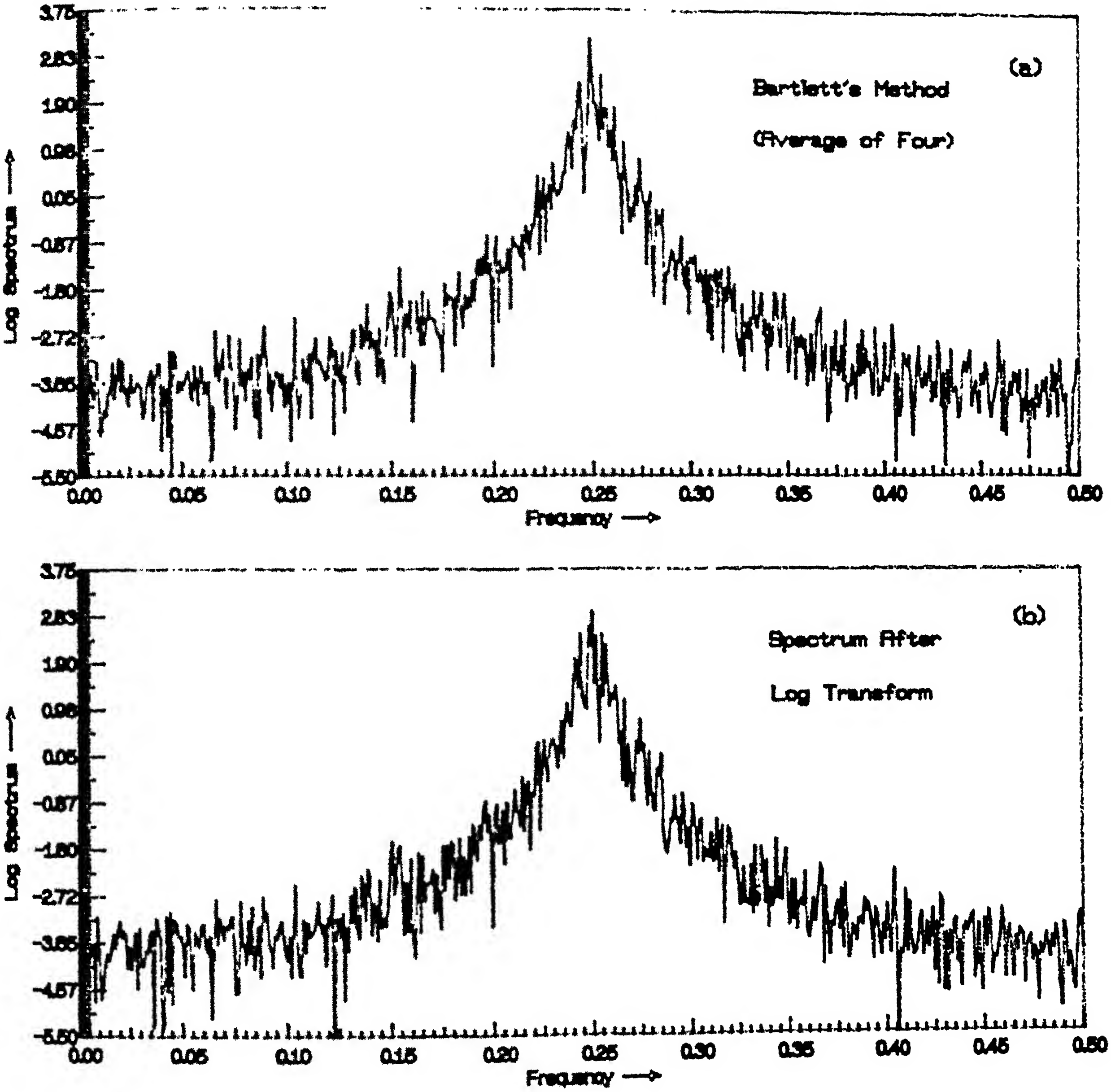


Figure 4.10: Spectrum Estimate of a Narrow Band Signal by Bartlett's Method. (Average of Four)

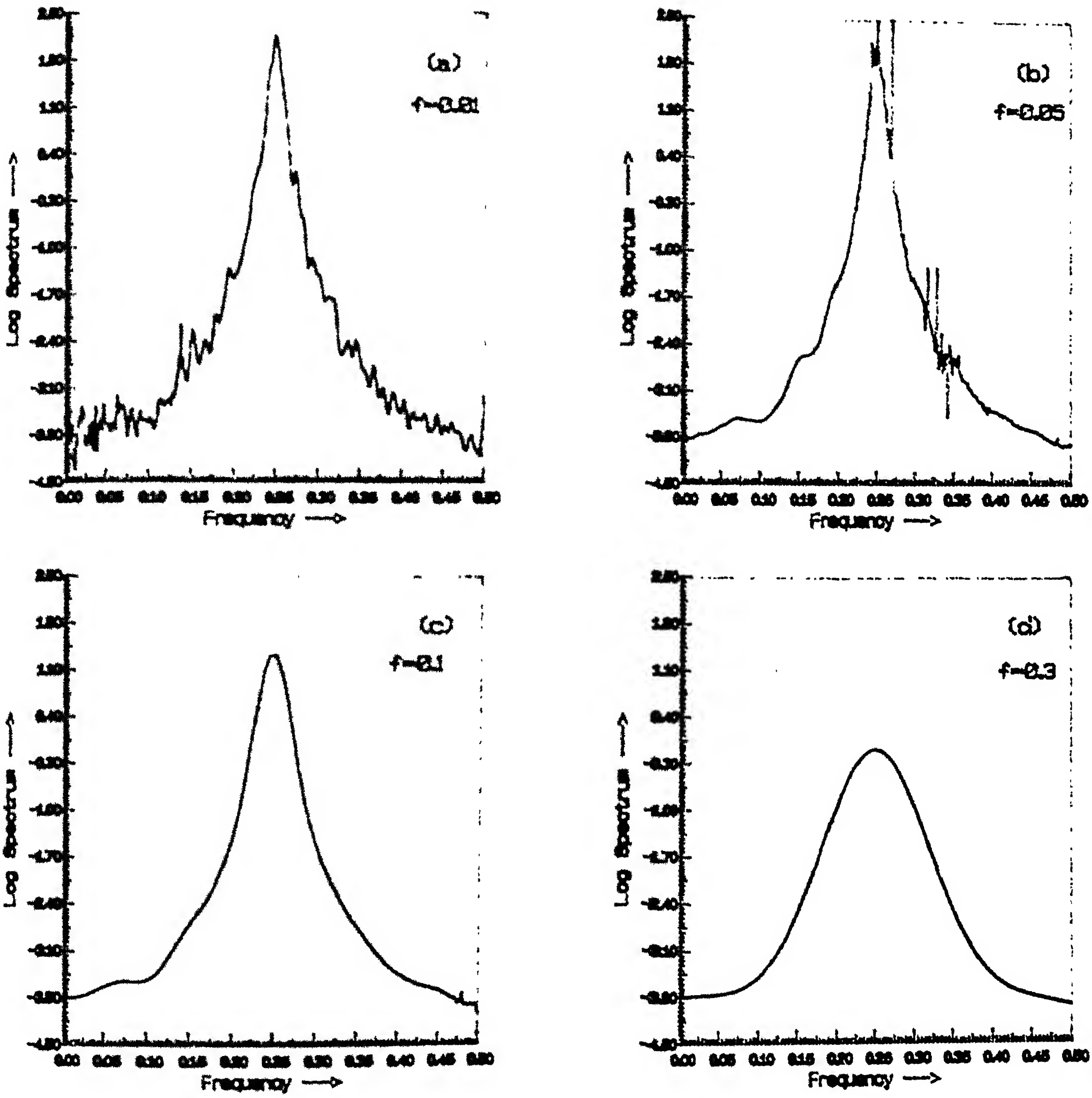
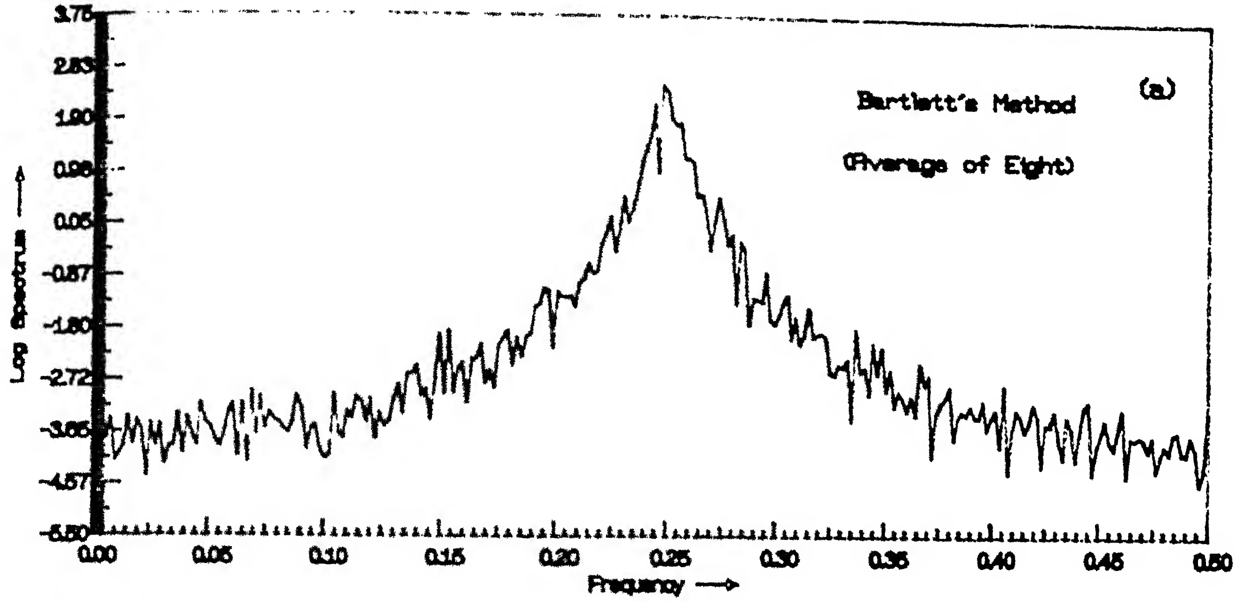


Figure 4.11: Spectrum Estimate of a Narrow Band Signal by ATS Method.(a=8)



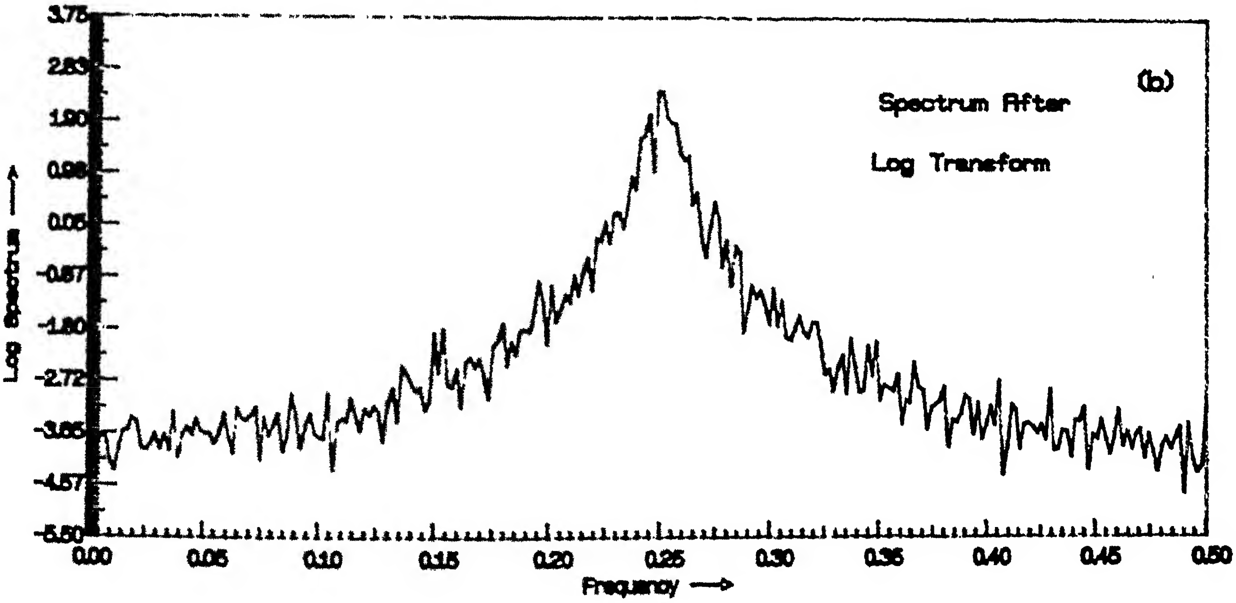


Figure 4.12: Spectrum Estimate of A Narrow Band Signal by Bartlett's Method. (Average of Eight)

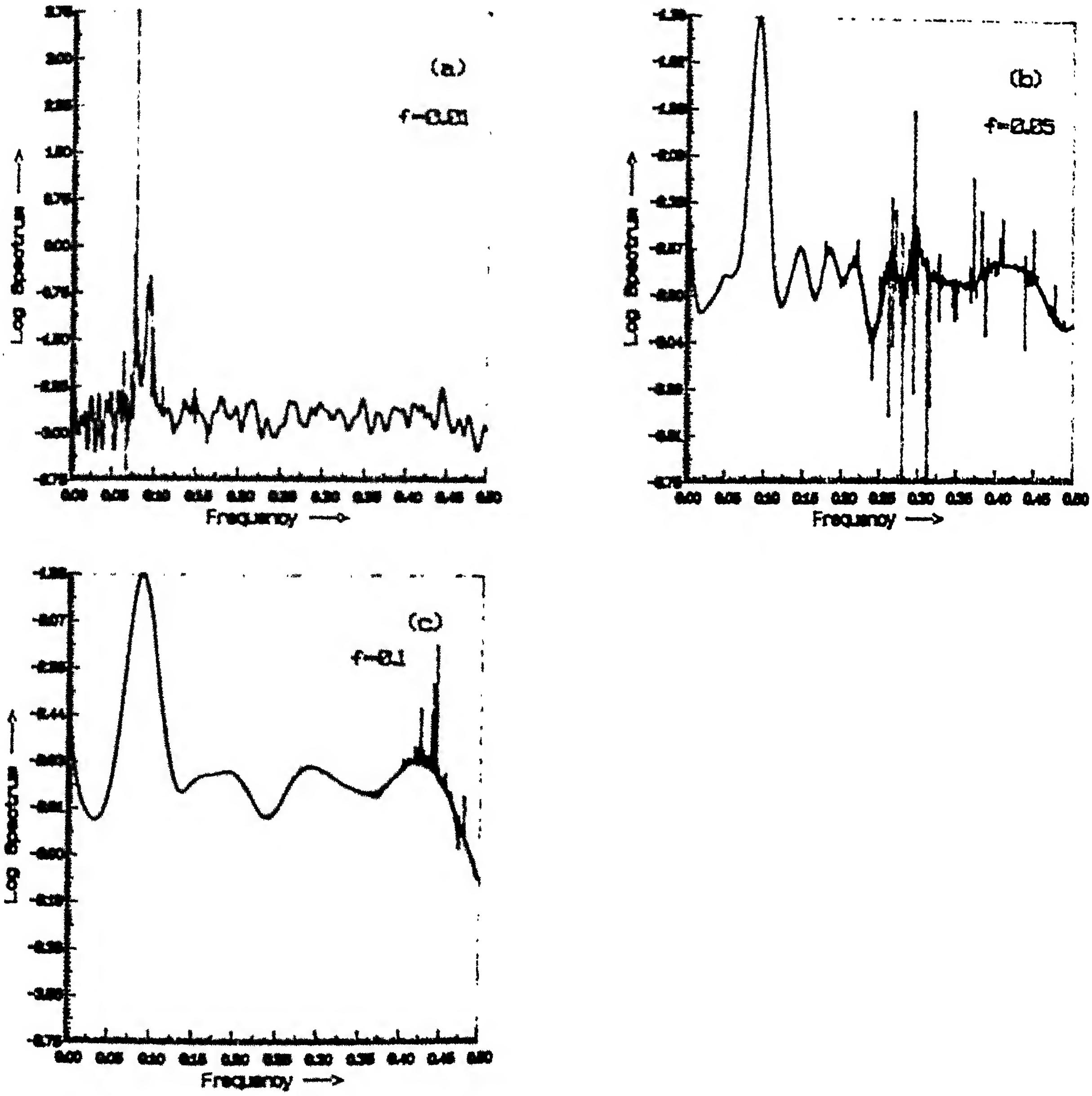


Figure 4.13: Spectrum Estimate of Sinusoids in Uniform White Noise by ATS Method.(a=2)

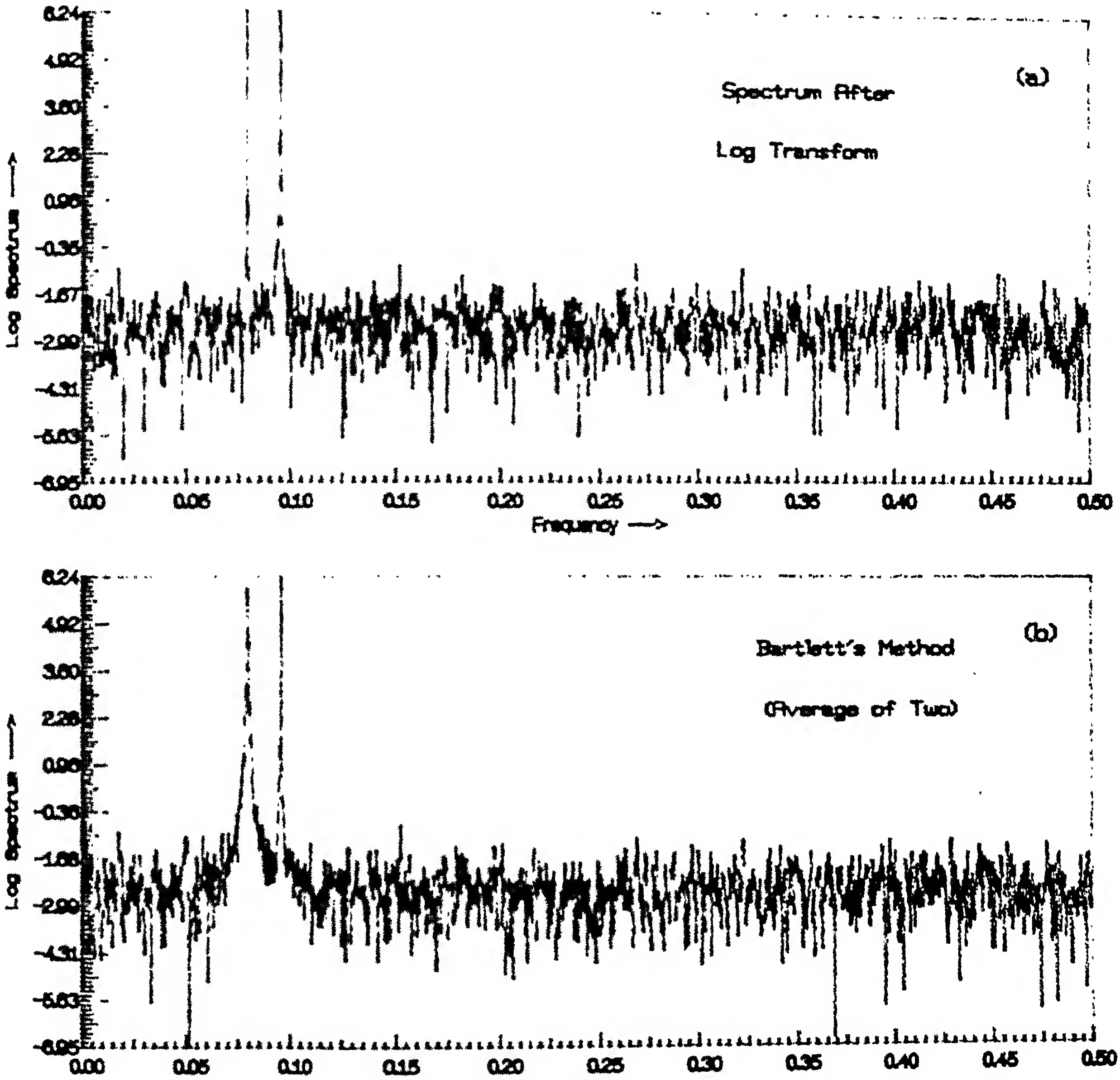


Figure 4.14: Spectrum Estimate of Sinusoids in Uniform White Noise by Bartlett's Method.(Average of Two)

Chapter 5

Future Scope

The ATS method essentially consists of estimating a regression surface(or curve) by smoothing. Much research has been carried out and there exists lot of literature in this field.

One important property of the ATS method (which has not been made use of in this work) is that the residuals obtained by regression can be used to determine thye adequacy of the fit. For this purpose the residuals are to be compared with the fitted values in the smoothed spectrum. Larger the smoothing factor larger will be the size of the residuals. Thus the residuals can be used as an important factor for deciding upon the amount of smoothing.

Another method to choose the optimum value of 'f; is to find it iteratively. The method is as described. Let $\hat{y}_i(f)$ be the fitted value at x_i for a given value of 'f' with y_i not included in the computations. The initial fitted value of $f = f_0$ is chosen by minimizing

$$\sum_{k=1}^{n}(y_k - \hat{y}_k(f))^2$$

where n is the number of points in the regression. Once this value f_0 is found out the robustness weights δ_k can be obtained as done in the loess algorithm. The next value of f is then chosen by minimizing

$$\sum_{k=1}^{n}\delta_k(y_k - \hat{y}_k(f))^2$$

This procedure is repeated till convergence takes place at the final value of f. This

technique has been used in some applications [6] although not for the periodogram.

The loess procedure does local fitting using the method of least squares. Another method is spline smoothing for the estimation of regression surface [19]. Here a smoothing procedure using splines for the log periodogram is discussed. Essentially we have looked into two aspects. One is the choice of the optimal smoothing factor 'f'. The other is an alternative to the locally weighted regression used in the ATS method.

Bibliography

[1] Abramowitz, M., and Stegun, I.A. (eds) (1964), *Handbook of Mathematical Functions,* Applied Mathematics Series No.55, Washington DC: National Bureau of Standards.

[2] Bartlett, M.S. (1947), *The Use of Transformations,* Biometrics Bulletin, Vol.3, No.1, pp 39-51.

[3] Bartlett, M.S., and Kendall, D.G. (1946), *The Statistical Analysis of the Variance and the Logarithmic Transformation,* Journal of the Royal Statistical Society Supplements, Vol.8, No.1, pp 128-138.

[4] Blackman, R.B., and Tukey, J.W. (1958), *The Measurement of Power Spectra,* Dover Publications Inc, New York.

[5] Bloomfield, P. (1976), *Fourier Analysis of Time Series: An Introduction,* John Wiley Publications.

[6] Cleveland, W.S. (1979), *Robust Locally Weighted Regression and Smoothing Scatterplots,* Journal of the American Statistical Association, Vol.74, No.368, pp 829-836.

[7] Cleveland, W.S., and Devlin, S.J. (1988), *Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting,* Journal of the American Statistical Association, Vol.83, No.403, pp 569-610.

[8] Cleveland, W.S., Mallows, C.l., and McRae, J.E. (1993), *ATS Methods: Non-Parametric Regression for Non-Gaussian Data,* Journal of the American Statistical Association, Vol.88, No.423, pp 821-835.

[9] Curtiss, J.H. (1943), *On Transformation used in the Analysis of Variance,* Annals of Mathematical Statistics, Vol.14, pp 107-122.

[10] Davenport, W.B. (1970), *Probability and Random Process,* MaGraw Hill Book Company, New York.

[11] David, F.N., Kendall, M.G., and Barton, D.E. (1966), *Symmetric Functions and Allied Tables,* Cambridge University Press.

[12] Freeman, M.F., and Tukey, J.W. (1950), *Transformations Related to the Angular and the Square Root,* Annals of Mathematical Statistics, Vol.21, pp 607-611.

[13] Jenkins, G.M., and Watts, D.G. (1968), *Spectral Analysis and its Applications,* Holden-Day Publications.

[14] Kay, S. (1988), *Modern Spectral Analysis,* Prentice Hall, New Jersey.

[15] Koopmans, L.H. (1974), *The Spectral Analysis of Time Series,* Academic Press.

[16] Oppenheim, A.V., and Schafer,R.W. (1975), *Digital Signal Processing,* Prentice Hall India Private Limited.

[17] Papoulis, A. (1990), *Probability, Random Variables and Stochastic Process,* MaGraw Hill International Edition.

[18] Press, W.H., Teulolsky, S.A., Vellering, W.T., and Flannery, B.P. (1988), *Numerical Recipes in C,* Cambridge University Press.

[19] Wahba, G. (1980), *Automatic Smoothing of the Log Periodogram,* Journal of the American Statistical Association, Vol.75, pp 122-132.